Quality & Quantity
https://doi.org/10.1007/s11135-024-01982-y

Cross-country analysis of science, technology and innovation policies: non-covid-19 related and Covid-19 specific STI policies in OECD countries

Margherita Russo[1] · Pasquale Pavone[2] · Dirk Meissner[3] · Fabrizio Alboni[4]

Accepted: 18 September 2024

Abstract
In OECD countries, Science, Technology and Innovation (STI) policies were seen as key aspects of coping with the Covid-19 pandemic. Now that the pandemic is over, identifying which policy mix portfolios characterised countries in terms of their non-Covid-19 related and Covid-19 specific STI policies fills a knowledge gap on changes in STI policies induced by exogenous shocks. The descriptive nature of this exercise sheds light on the emergency phase, which was addressed in different ways by countries with similar STI policy portfolios in the last decade before the pandemic. Using information on STI policy initiatives in OECD countries, this paper proposes a multidimensional analysis to classify policy initiatives based on both codes (of innovation policy themes, policy instruments and target beneficiaries) and free text policies' descriptions. Based on text mining and clustering techniques, the multidimensional analysis highlights semantic similarities between the combinations of codes and terms, making it possible to identify policy mixes that characterise non-Covid-19 related and Covid-19 specific STI policies. The cross-country comparison draws attention to the specific policy mix portfolios implemented by countries during the pandemic. The paper contributes to the literature on innovation policy mix in terms of research methods and results in identifying STI policy portfolios and groups of countries with similar structural composition of their innovation policy portfolios, implementing a range of STI strategies in tackling the pandemic. Policy implications of the findings are discussed, with a forward-looking perspective for the analysis of post-pandemic STI policies.

Keywords STI policies · Cross-country analysis of STI policies, STI policy mix · Covid-19 · Multidimensional analysis · Natural language processing · Semantic analysis · OECD-EC STIP Compass

Extended author information available on the last page of the article

Published online: 25 September 2024

1 Which STI policy portfolio did countries adopt during the pandemic?

A country's Science, Technology and Innovation (STI) policy is embedded in the country's general policy agenda which is unique to individual countries. STI policy portfolios are horizontal, with policy measures touching on diverse policy fields, including science, health, education, employment, migration, trade, taxation, infrastructure, investment, SMEs and competition, to name a few. A country's STI policy measures follow two major streams: the structural and systemic stream, which stresses the links between actors, and the economic and financial (both budgetary and private) stream, targeting the allocation of funds and encouraging business innovation (Gokhberg et al. 2022). For different reasons, countries might reshape their STI policy mix to align with a new agenda or they might be forced to revise the mix due to exogenous shocks (Meissner and Kergroach 2021), as in the case of the Covid-19 pandemic. Uncertainties during the Covid-19 pandemic (Boin and Lodge 2021) pushed countries to design and implement policy initiatives, in many domains strongly related, to coping with urgent interventions: from economics to health, research and education (Gershman et al. 2022). These interventions forced countries to establish new STI policy initiatives or to revise the scope and focus of existing ones. During the Covid-19 pandemic, it also became obvious that policy makers had clear expectations that science would be able to address the challenges of the pandemic within a tight timeframe. It was frequently understood that science would deliver models of the pandemic's development and measures considered reasonable to limit the pandemic's impact, to protect people and the like. In this context, it appeared that science reacted to these requirements and challenges by undertaking reasonable efforts to make a contribution to fighting the pandemic, but the proposed solutions were often only at an exploratory stage rather than fully based on evidence (Dziembała 2023; Yun 2023). On the other hand, policies took advantage of scientific advice to justify policy measures which had an impact on people in one way or another. In addition, science has become more of an issue of public debate and discussion in the media, with scientists more often and frequently involved in policy discussions (Bozeman 2022).

STI policy literature addresses the rationale of stringent national restrictions (Plümper and Neumayer 2022) about the government narratives on the economic cost of the pandemic falling on selected groups (Narlikar and Sottilotta 2021; Dziembała 2023), on school closures (Toshkov et al. 2021), and also concerning aspects of decentralised administration (Capano and Lippi 2021; Toshkov et al. 2021). The emergency policy per se has been the subject of several contributions (Schmidt 2021). A specific strand of literature, with a focus on science policy as a lever for coping with pandemic challenges, has been addressed by authors who highlight the relevant dimensions of STI policies and a set of tools (patents, prizes, subsidies) to support science development (Abi Younes et al. 2021) and the challenges for enhancing proactive STI policies (Baute and Ruijter 2021). Systemic state interventions, speeding up and scaling up solutions provided by science and technology research and citizen engagement, and cooperation across many diverse domains were considered key dimensions to be leveraged in reducing the impact of the Covid-19 pandemic, within and across countries, in particular in the OECD recommendation on STI policies (OECD 2020, 2023).

In this paper, we review which STI policies were implemented by countries in the first two years of the pandemic. Our focus, in particular, is on additional policy initiatives not yet implemented in the countries' agendas which were specifically introduced to address Covid-

19. We adopt a multidimensional analysis of STI policies in OECD countries to answer three research questions:

1) Which policy mixes characterise the STI policies implemented by countries as their structural STI policies and about those implemented to cope with the pandemic?
2) Which policy mix portfolios characterise non-Covid-19 related and Covid-19 specific STI policies in OECD countries?
3) Did countries with similar structural STI policy portfolios adopt similar or different Covid-19 specific policy portfolios?

To answer our research questions, we created an ad hoc dataset, drawing on OECD sources, and we performed a multidimensional analysis to classify policy initiatives based on both codes (of STI policies' thematic areas, policy instruments and target beneficiaries) and free texts of policies' descriptions. A classification of policy mixes can be implemented by using various models of topic detection (Antons et al. 2020; Alboni et al. 2023). In this paper, we adopt, for each set of policy initiatives, first, a text mining strategy to select terminology in the free text description of policy initiatives. We then develop a vector space model and apply a clustering technique to the set of selected terms and multiple codes of policy initiatives. The multidimensional analysis highlights semantic similarities between the combinations of codes and terms, making it possible to identify clusters of policy initiatives that characterise non-Covid-19 related and Covid-19 specific STI policies. The resulting clusters are interpreted as policy mixes. For each country, the policy portfolio is then defined, i.e. the policy mixes which are characteristic (in statistical terms) of that country. The cross-country comparison of policy portfolios is interpreted concerning non-Covid-19 related and Covid-19 specific innovation policies, pointing out their structural similarity, i.e. the dimensions characterising their policy mix.

The paper contributes to the literature on the STI policy mix in terms of multidimensional methods and with the identification of STI policy portfolios and groups of countries that with similar structural compositions of their innovation policy portfolios adopted a range of strategies in addressing the pandemic with STI policy initiatives. The descriptive nature of this exercise sheds light on the emergency phase that has been addressed in different ways by countries that had similar structural STI policy portfolios.

In what follows, we present the data available from the OECD *STIP Compass* and the OECD *STIP Covid-19 Watch* policy initiatives (Sect. 2) and the rationale for analysing each dataset separately. Section 3 outlines the methodology adopted for multidimensional analyses to identify the policy mixes that characterise the STI policy initiatives and to identify which groups of countries have similar policy portfolios. The results are discussed in Sect. 4. Section 5 highlights the limitations and strengths of the methodology adopted and the dataset created for our analysis, Sect. 6 concludes with a summary of results, and Sect. 7 discusses the policy implications of the methods and results, to broaden the scope of the analysis to post-pandemic STI policies. Supplementary figures and tables complement the detailed results and analyses. Results can be browsed online by using the Tableau Public navigation tool available at https://public.tableau.com/app/profile/pasquale.pavone/viz/CrossCountryAnalysis/STIPolicies.

2 Data sources

We rely on two datasets made available by the OECD (2024), the most complete source of information entered by the respective National Contact Points (NCP) on STI policies in the OECD countries: the STIP Compass (2021 edition) and the STIP Covid-19 Watch, downloaded on 19 February 2022.

The elementary unit of our analysis – a policy "document" is the policy initiative, with information encompassing a free text describing each policy initiative (objectives and description) and a set of codes to classify them concerning 59 categories of policy themes, 28 policy instruments and 31 target groups of beneficiaries (EC-OECD 2021; Russo and Pavone 2021). In the STIP Covid-19 Watch, there are also 28 new codes for policy themes, created by the OECD to specify themes related to the pandemic (for details on the datasets and the codes, see Supplementary Table S1).

The STIP Compass dataset, comprising information on 6,285 policy initiatives, refers to 62 territorial entities (countries, regions and the European Union). Except for Belgian regions, the STIP Compass had no information on policy initiatives at a sub-national level. Start dates range from 1900 to 2021, with 76.72% of policy initiatives being issued in the years 2010–2021 and 15.15% in the years 2020–2021 (Supplementary Table S2).

During the emergency phase of the pandemic, in the OECD countries, about half of the STI policy initiatives were specific to Covid-19, and were classified in the STIP Covid-19 Watch dataset. This dataset encompasses 935 STI policy initiatives that refer to 61 territorial entities[1]. Start dates range from 1952 to 2021, with 88.67% of policy initiatives being issued in the years 2020–2021 with 80.43% in 2020 alone (see Supplementary Table S2), the year in which the World Health Organization identified the pandemic (WHO 2024)[2]. Covid-19 policy initiatives initiated before 2020 and that are included in that dataset are assumed to be structural STI policies that countries relied on to cope with the pandemic, such as the one from Switzerland "to support basic research (bottom-up)" (started in 1952), from Norway "to develop ideas and start-up environments, to promote firm growth and support competitive innovation clusters in their internationalisation processes" (started in 1968), and from the Russian Federation "to create a stable and secure information and telecommunications infrastructure for high-speed transmission, processing and storage of large amounts of data, accessible to all organizations and households" (started in 2018).

In addition to the obvious differences between the two sets of policies, in terms of the specific lexicon and ad hoc policy themes adopted in the Covid-19 specific policy initiatives, we observe a sharp characterisation in terms of policy instruments and target beneficiaries, as shown in Fig. 1, concerning the significant standardised chi-squared residuals of the distribution of codes in the two datasets (values greater than 2). About STIP Covid-19 Watch policy initiatives (Fig. 1, panel a), the policy instruments encompass, essentially, those specific ones because of the role attributed to science in coping with the pandemic (i.e. the formal consultation of stakeholders or experts, procurement programmes for R&D and innovation, project grants for public research, information services and access to datasets, institutional funding for public research, public awareness campaigns and other outreach activities, grants for business R&D and innovation, the creation or reform of governance

[1] The updated edition of the STIP Covid-19 Watch (April 2023) differs for only one policy initiative.

[2] WHO announced the end of the emergency phase of COVID-19 in May 2023 https://www.who.int/newsroom/fact-sheets/detail/coronavirus-disease-(covid-19).

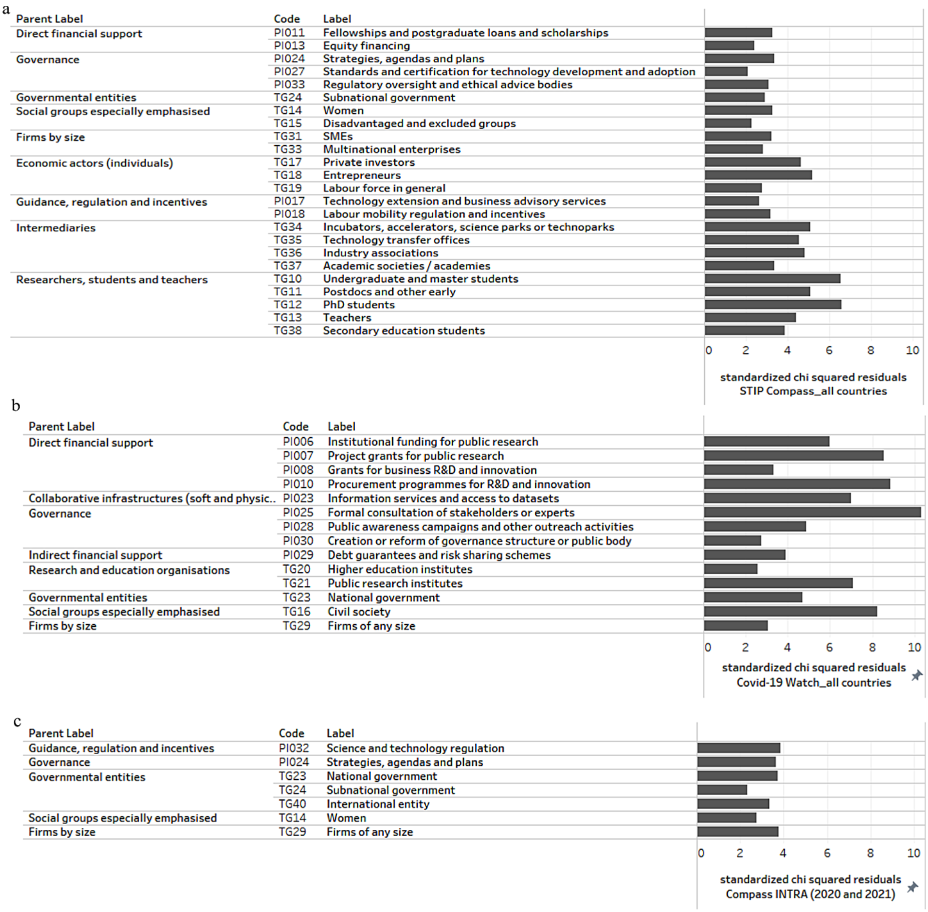


Fig. 1 STIP Compass and STIP Covid-19 Watch: distribution of standardised chi-squared residuals by type of policy instrument (PI) and target group (TG) of the STI policy initiatives Codes about PIs and TGs: in the comparison of STIP Compass and STIP Covid-19 Watch (panels a and b), and in the comparison of STIP Compass pre-2020 and in the period 2020–2021 (panel c). (**a**). Focus on STIP Covid-19 Watch, in comparison with STIP Compass. (**b**). Focus on STIP Compass, in comparison with STIP Covid-19 Watch. (**c**). STIP Compass: focus on 2020–2021, in comparison with pre-2020 policy initiatives. Source: authors' analysis of STIP Compass (2021 edition) and STIP Covid-19 Watch, download 19.02.2022

structures or public bodies); the target groups are, primarily, broad categories of beneficiaries in society, public and private sectors (specifically, civil society, public research institutes and higher education institutes, national governmental entities, firms of any size). When comparing the distribution of categories of policy instruments and target groups of beneficiaries, the STIP Compass (Fig. 1, panel b) shows a variety of other specific dimensions that are typical of those policies (Russo and Pavone 2021). In general, the policy instruments and target groups that characterise the policy initiatives included in the STIP Compass vary significantly throughout the whole time frame analysed (details in Supplementary Table S3). However, if we compare the policy instruments and target groups of the STIP Compass before (until 2019) and during the pandemic (2020 and 2021), with the results in Fig. 1, panel c, we observe some specific dimensions similar to those of the Covid-19 Watch STI policy initiatives. They highlight a relatively higher importance of policy instruments focusing on science and technology regulation and the involvement of governmental entities at international, national and also sub-national levels, and a target on relatively more vulnerable beneficiaries in the pandemic, such as firms of all sizes and women. These elements are not, however, sufficient to answer the research questions addressed in this paper, which require a multidimensional analysis - presented in Sect. 4 - to reveal the rich characterisation of the two sets of policy initiatives in terms of their policy mix, relevant for the cross-country comparison of their policy portfolios.

3 Methodology and data analysis

The various steps to prepare and analyse the data needed to implement the cross-country comparison (including also regions and the European Union) of non-Covid-19 related and Covid-19 specific STI policies can be described in the four main steps outlined in Fig. 2 (details on cleaning the data inputs are available in Supplementary Material Annex 1).

Given the specific characteristics of the two datasets in question, they were analysed separately but using the same methods. Specifically, the policy initiatives relating to the Covid-19 Watch have a highly specific lexicon that is notably different from that used before the pandemic, even for policies in a healthcare context, and new themes have been created for coding such policies. Therefore, combining the two datasets into one would have made it difficult to achieve an interesting level of detail regarding intra-Covid-19 policies. Nonetheless, an attempt at a combined analysis was made, confirming the hypothesis of a simplistic characterisation of those policies. These results are available upon request.

For each dataset, three steps encompass, respectively: Step 1, the preparation of a lexical textual analysis to extract the content terms of the free text descriptions and the objectives of each policy initiative, which will be used together with multiple codes of policy initiatives to define a vector space model; Step2, the identification of policy mix; Step 3, the identification of cross-country similarities in policy portfolios, i.e. the policy mix characteristic of a country. The cross-country comparison of policy mix portfolios is interpreted concerning non-Covid-19 related and Covid-19 specific innovation policies. A fourth step compares non-Covid-19 related and Covid-19 specific STI innovation policy portfolios

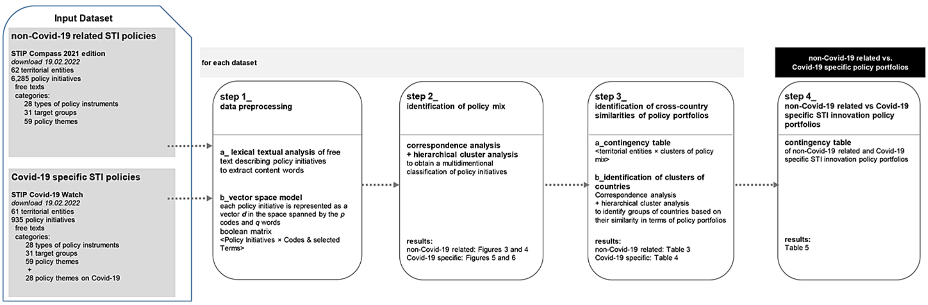


Fig. 2 Workflow for the cross-country comparison (including regions and the European Union) of non-Covid-19 related and Covid-19 specific STI innovation policies. Source: authors' analysis

to shed light on similarities/differences between countries that had similar structural STI policy portfolios.

Below, we describe the methods used in each step of the analysis of the policy initiative datasets.

3.1 Step 1: data pre-processing

3.1.1 Step 1a: lexical textual analysis

In this step, we deal with the identification and extraction of the content terms present in the free text part used to describe each policy initiative together with its objectives.

For this purpose, we define a corpus for each dataset, understood as a collection of texts to be analysed from a lexical-textual point of view. Considering that not all the words in a text have semantic content, we define a text mining procedure to select a reduced number of terms which can represent the semantic content expressed in the description of the policy initiatives. A text mining strategy, based on the software package TaLTaC[3] (Bolasco et al. 2016), was implemented to select the content terms. To this end, for each corpus, the following procedure was carried out: (a) the words were lemmatised and grammatically annotated; (b) nominal multi-word expressions were recognised (Pavone 2018); (c) the content terms with at least 5 occurrences, grammatically annotated as nouns and adjectives, were selected.

The STIP Compass dataset consists of 6,285 policy initiatives. The free text of the policy objectives and description is not available for 51 policy initiatives. In addition, 10 policy initiatives have a description text in Spanish, and then those texts cannot be used in the lexical analysis of the corpus in English, while another 19 policies refer to bilateral agreements between Brazil and other countries, effectively reporting the same text repeated with very few changes in each case. Both groups of texts were excluded from the free text analysis. These 80 policies (without free texts, with texts in Spanish and with repeated texts) are analysed only through the codes that have been defined for them. Therefore, the STIP Compass corpus - created to identify the selected terms consists of 6,205 texts and results composed of a vocabulary of 22,818 different words (types) with an overall number of 496,073 occurrences (tokens). The text mining procedure applied on the STIP Compass corpus has allowed us to select 3,321 terms, of which 825 are multiword expressions. Specifically, the following were selected: 2,496 lemmas classified as nouns or adjectives, with at least 5 occurrences in the text and present in at least 5 policy initiatives; 825 multiword expressions with at least 10 occurrences in the corpus and present in at least 5 policy initiatives. The following are the 30 most occurring multiword expressions: *higher education*, *research institution*, *private sector*, *scientific research*, *research project*, *public sector*, *artificial intelligence*, *research infrastructure*, *technology transfer*, *technological development*, *young people*, *human resource*, *new technology*, *intellectual property*, *climate change*, *research result*, *applied research*, *research institute*, *basic research*, *economic growth*, *financial support*, *young researcher*, *international cooperation*, *research organisation*, *innovation policy*, *digital transformation*, *gender equality*, *innovative solution*, *sustainable development*, *public policy*.

[3] TaLTaC is the acronym for "Trattamento automatico Lessico-Testuale per l'analisi del Contenuto" (lexical-textual automatic treatment for content analysis). It has been designed for automatic text analysis in the dual logic of Text Analysis and Text Mining (Bolasco et al. 2016).

In the case of the STIP Covid-19 Watch dataset, there are two policy initiatives without a free text description. Therefore, the corpus under analysis consists of 933 policy initiatives and results composed of a vocabulary of 6,889 different words (types) with an overall number of 57,212 occurrences (tokens). Specifically, 802 lemmas and 25 multiword expressions were selected. The following are the 25 multiword expressions identified: *research project*, *infectious disease*, *public health*, *clinical trial*, *research institution*, *scientific advice*, *innovative solution*, *higher education*, *private sector*, *support research*, *scientific community*, *health system*, *scientific research*, *financial support*, *social science*, *innovation project*, *research institute*, *scientific knowledge*, *health care*, *supply chain*, *genome sequencing*, *medical device*, *data sharing*, *research activities*, *health emergency*.

After the identification of the terms under analysis in the two corpora, we can summarise their main characteristics (Table 1): the size of the two corpora and the variety of terms are adequate for automatic text analyses. The cross-country comparison of data in terms of the share of occurrences of terms and the share of policy initiatives shows that some countries present a lower share of occurrences of terms in the STIP Covid-19 Watch (such as Poland, Austria, Colombia, Hungary, Australia), while other countries and territorial entities have a significantly higher share of occurrences (Canada, Greece, the European Union, the Russian Federation, Czech Republic…). Apart from a few cases, the share of occurrences of selected terms is largely related to the number of policy initiatives (detailed figures in Figures S1-S3).

Table 1 Main characteristics of the two corpora in text analysis

	STIP Compass	STIP Covid-19 Watch
no. of policy initiatives	6,285	935
no. of policy initiatives with no textual description & objectives	51	2
no. of policy initiatives with a textual description & objectives	6,234	933
no. of policy initiatives with non-English text (Spanish)	10	0
no. of policy initiatives excluded due to redundancy problems	19	0
no. of text descriptions & objectives composing the corpus under analysis	6,205	933
Types	22,818	6,899
Tokens	496,073	57,212
Length of free text descriptions & objectives: tokens, min-max	min 4 - max 4,634	min 8 - max 488
% of policy initiatives with fewer than 50 occurrences	30.9	47.2
% of policy initiatives with more than 100 occurrences	22.2	11.8
Mean of token by policy document	79.9	61.2
Median of token by policy document	66	51
Types/Token ratio	0.05	0.12
hapax	11,198	3,339
% hapax by types	491	48.4
% hapax by token	2.2	5.8
selected terms	3,321	827
selected lemmas > 5 occurrences	2,496	802
identified and selected multiword expressions > 10 occurrences	825	25

Source: authors' analysis of STIP Compass (2021 edition) and STIP Covid-19 Watch, download 19.02.2022

3.1.2 Step 1b: Vector space model representation

Once the content terms of each policy initiative have been selected, we can define each data-set to be analysed using a vector space model (Salton et al. 1975) in which each policy initiative, i, can be represented as a vector $\mathbf{d_i}$ in the space spanned by the p codes and q words:

$$\mathbf{d_i} = (c_{i1}, \ldots, c_{ip}, w_{i1}, \ldots, w_{iq})$$

In this operation we are considering codes and words in the same way. In fact, just as with words, each code has a semantic value represented by its definition (for more detail see Supplementary Table S1). A given code is mentioned only once in a document, therefore we will also weight the selected words, not according to their frequency, but exclusively according to their presence/absence in the documents, assigning 1 or 0 to codes and words as they appear or not in each document.

The codes encompassed in the matrix are only those used in coding, respectively, 107 (out of 118) and 126 (out of 146), for STIP Compass and STIP Covid-19 Watch; the words encompassed in the matrix are the selected terms identified in the lexical textual model (summarised in Table 1). The vector space model representation of policy initiatives, applied to codes and terms, has generated two Boolean matrixes, respectively ⟨*6,285 STIP Compass Policy Initiatives* × *107 Codes&3,321 selected Terms*⟩ and ⟨*935 STIP Covid-19 Watch Policy Initiatives* × *126 Codes&827 selected Terms*⟩.

The creation of a vector space model to describe the policy initiatives in terms of both codes and selected terms is the main difference concerning the method proposed by Russo and Pavone (2021) in the analysis of the STIP Compass dataset (2019 edition), who used separate analyses of codes and selected terms and then combined the results.

3.2 Step 2: identification of policy mix

For each dataset, the Boolean matrix is analysed with a correspondence analysis using the R Package *FactoMineR* (Lê et al. 2008) to obtain the dimensionality reduction of the matrix. A hierarchical cluster analysis on the results of the correspondence analysis (first 10 factors) has been implemented using the R package *NbClust* (Charrad et al. 2014), employing Euclidean distance and Ward's method (Ward 1963) to classify the policy initiatives concerning both codes and terms. Each cluster returns groups of policy initiatives that are similar to each other in terms of their contents. This homogeneity makes a semantic character prevailing in each cluster explicit. The interpretation of this semantic character allows us to label each group of policy initiatives according to their mix of policy themes, instruments, target groups of beneficiaries, and specific terms describing the policies. In particular, this interpretation is done by reading the dictionaries of codes and terms, ranked by their test value. Analogous to the value of a standardised normal variable (under the hypothesis of independence), the test value lies between the values –1.96 and +1.96 (with a probability of 0.95) (Lebart et al. 1998, p. 98 and p. 137). For each cluster, the list of distinctive codes and terms has been interpreted to label the policy mix characteristic of each cluster.

Through factorial maps, it will be possible to graphically observe the results of the correspondence analysis and cluster analysis. A factorial map displays the relationships between the rows and columns of the matrix under analysis in a low-dimensional space. Rows and columns that are similar in terms of their associations with other rows and columns tend to cluster together. The distance between points in the map reflects the strength of association or dissimilarity between the corresponding rows or columns in the contingency table. In both analyses, the factorial map representing the combination of the first two factors *f1f2* will be displayed in the text, while the factorial maps of the first ten factors are in Supplementary materials.

In the identification of the STI policy mix, the statistical method outlined above is the same proposed by Russo and Pavone (2021), while Howoldt (2024) implements a probabilistic model. In both papers, the authors rely on data from the OECD STIP Lab, although Howoldt (2024) focuses only on a selection of input data. In this paper, the rationale for our choice is that correspondence analysis to obtain the dimensionality reduction of the matrix - together with hierarchical clustering on factors allow greater control in interpreting results, in particular in the decision regarding the level of detail in grouping policy initiatives that is relevant for the analysis (an issue discussed by Alboni et al. 2023, in a comparative analysis of some of the most widely used methods of topic detection used in the literature).

3.3 Step 3: identification of cross-country similarities of policy portfolios

After having obtained clusters of STI policy initiatives, we proceed to the identification of groups of countries based on their similarity in terms of their STI policy mix. We create a matrix ⟨*territorial entities* × *clusters of policy mix*⟩ (step 3a) where the weights correspond to those used in the previous matrix, i.e. the sum by country of the presence/absence of the codes and terms within the policy initiatives. On this matrix, we perform a correspondence analysis and a hierarchical cluster analysis on the first ten factors (step 3b). Through the latter, we obtain groups of countries according to their structural similarities in terms of their policy mix portfolios. While the similarity between policy initiatives is of a semantic type, in the context of cross-country comparison, the similarity between countries is of a structural type.

3.4 Step 4: comparison of non-Covid19 related and Covid19-specific STI innovation policy portfolios

By combining the results obtained in step 3, it is possible to construct a contingency table of countries' non-Covid-19 related and Covid-19 specific STI innovation policy portfolios. This allows the identification of cross-country similarities.

4 Results

Below we look in detail at the results of steps 2, 3 and 4 of our analysis. Both for the OECD STIP Compass and for the OECD STIP Covid-19 Watch, the resulting groupings constitute policy mixes of policy initiatives. These policy mixes are labelled according to their characterisation based on codes and terms. To distinguish the results of the two matrices, the policy mix labels obtained in the STIP Compass analysis are preceded by the letter "c" followed by an increasing number, while the policy mix labels obtained in the STIP Covid-19 Watch analysis are preceded by the letters "cl-" followed by an increasing number.

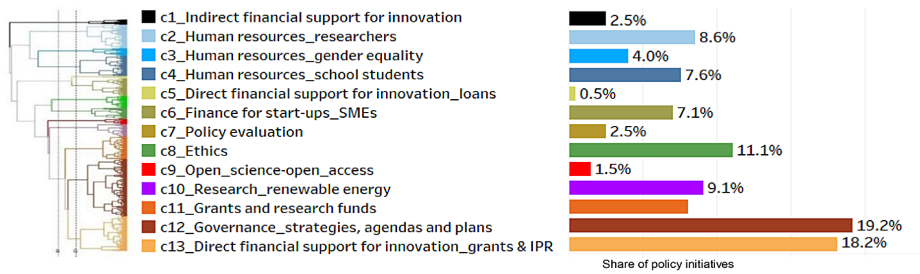
Concerning each dataset of STI policies, in Sect. 4.1 and 4.2 we describe which STI policy mixes can be identified and the corresponding countries' policy portfolios. In Sect. 4.3 we present the cross-country comparison of non-Covid-19 related and Covid-19 specific policy portfolios.

4.1 STIP compass

4.1.1 STI policy mix

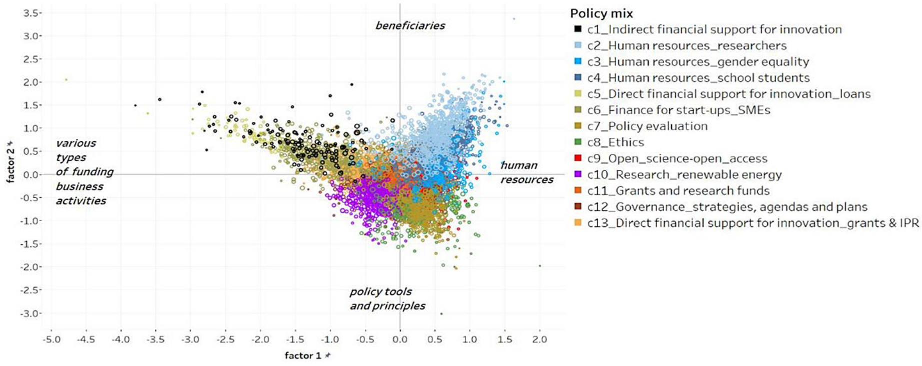
Based on their semantic similarity of codes and terms, we identify 13 clusters of policy mixes in the STIP Compass. According to the Caliński-Harabasz index (Caliński and Harabasz 1974, see Supplementary Figure S4), the optimal number of clusters is ten, but we preferred the thirteen cluster cut (see Fig. 3), which allows us to split a large area of interventions into three different groups referring to grants and research funds (c11), strategies, agendas and plans (c12), and direct financial support for innovation (c13). The graphical results of clustering on codes and terms are shown in the factorial map *f1f2* in Fig. 4 (while Supplementary Figure S5 shows the factorial maps obtained from all combinations of the first 10 factors under analysis). Each dot in the factorial map is a policy initiatives document, coloured according to the policy mix cluster in which the document has been classified by the clustering algorithm. Cluster labels were assigned by reading the characteristic codes and terms in each cluster (listed in Table S4 in decreasing order of their test value). As an example of such a labelling exercise, let us consider c8, "Ethics" (quotations from characteristic terms and codes are in italics):

> "Ethics" focuses on the issues of algorithm development and artificial intelligence and has at its core integrity, trust and responsibility, research misconduct, data protection and privacy. Characteristics of these innovation policies are a set of policy instruments of the regulatory framework (Guidance, regulation and incentives, PI032) and Governance through Regulatory oversight and ethical advice bodies (PI033), Formal consultation of stakeholders or experts (PI025), Standards and certification for technology development and adoption (PI027), Creation or reform of governance structure or public body (PI030), Policy intelligence (e.g. evaluations, benchmarking and



Source: authors' analysis of STIP Compass (2021 edition), download 19.02.2022

Fig. 3 STIP Compass - Dendrogram of the hierarchical cluster of policy initiatives, classification based on multiple coding and terms (on the left), legend of colours and labels of the clusters (centre) and share of policy initiatives by cluster (histogram on the right). Share of policy initiatives. Source: authors' analysis of STIP Compass (2021 edition), download 19.02.2022



Source: authors' analysis of STIP Compass (2021 edition), download 19.02.2022

Fig. 4 STIP Compass - Factorial map *f1f2* - Distribution of Policy Initiatives – Matrix (*6,285 Policy Initiatives* × *107 Codes&3,321 selected Terms*) (2,685×3,428**).** Each dot corresponds to a policy initiative, with size proportional to the number of codes and terms. Dot colour is by policy mix detected by clustering. Source: authors' analysis of STIP Compass (2021 edition), download 19.02.2022

forecasts) (PI031). These instruments are about Research and innovation for society, specifically on Ethics of emerging technologies (TH89), and Public research systems, concerning Research integrity and reproducibility (TH21). An array of target groups is involved in such policies: Civil society (TG16), National government (TG23), International entity (TG40), Multinational enterprises (TG33), Private research and development lab (TG22), Labour force in general (TG19), but also Private investors (TG17).

In the factorial map *f1f2* presented in Fig. 4 there is a marked polarisation of policy mixes along the two factors (see Supplementary Figure S5 to interpret the polarisation), with a barycentre policy mix in which the governance orients interventions of national and sub-national governments on the innovation agendas of private and public agents (c12), the first factor characterised with respect to resources to support innovation (monetary and human ones) and the second factor to support innovation implementation (which policies and which beneficiaries).

In particular, along the first factor (from left to right), the polarisation unfolds from funding innovation in firms (indirect and direct financial support, respectively c1, c5 and c13) and finance for start-ups (c6), with a core theme on research into renewable energy (c10), vs. human resources (researcher, gender equality, and students, respectively c2, c3 and c4), with grants and research funds (c-11), and policy tools and visions: policy evaluation, ethics and open science (respectively, c7, c8, c9). Along the second factor, the polarisation (from bottom to top) is between policy tools and principles (evaluation, ethics, renewable energy, open science, respectively, c8, c7, c10, and c9), oriented by the national government (c12) towards an array of distinct beneficiaries (private companies, c1 and public researchers, c2 and c4).

The 13 policy mixes are not evenly distributed over time. With regard to the significant standardised chi-squared residuals of the distribution of the policy mixes by year (Table 2), it is possible to observe some brand new areas of intervention, characterising the most recent

Policy mix	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	Share of policy initiatives
c1_Indirect financial support for innovation	2.22	1.71	0.06	1.15	0.10	0.72	0.01	1.45	-1.70	-1.93	-0.07	-1.78	2.5%
c2_Human resources_researchers	1.48	0.81	0.37	2.16	1.72	-1.79	0.14	2.61	0.65	-1.81	-3.51	-1.31	8.6%
c3_Human resources_gender equality	-1.61	0.83	-0.24	-0.77	0.09	-0.94	0.65	-1.32	-0.63	-0.07	0.51	2.88	4.0%
c4_Human resources_school students	1.20	1.76	-0.85	-0.22	-0.62	0.93	0.56	0.07	0.25	-2.18	-0.20	0.37	7.6%
c5_Direct financial support for innovation_loans	-1.33	-0.03	0.92	2.34	0.48	1.34	0.19	0.75	-0.40	-0.56	-2.31	-0.96	0.5%
c6_Finance for start-ups_SMEs	-0.23	-0.41	3.03	-0.72	1.08	0.59	1.85	-0.50	0.05	-0.12	-2.21	-1.63	7.1%
c7_Policy evaluation	0.09	0.13	1.64	-0.24	0.92	1.85	-1.31	-1.02	1.28	-0.90	-1.02	-0.44	2.5%
c8_Ethics	-2.81	-1.27	-1.94	-3.27	-2.34	-2.01	-2.79	-1.21	2.12	6.67	6.31	-1.47	11.1%
c9_Open_science-open_access	1.24	-0.72	2.36	2.14	0.02	0.06	0.30	0.38	-0.20	-1.40	-1.13	-1.61	1.5%
c10_Research_renewable energy	-1.86	-2.96	-1.70	-3.06	-2.94	-2.46	-1.13	-2.35	-1.00	-1.13	9.96	7.27	9.1%
c11_Grants and research funds	1.26	1.44	1.24	1.36	0.71	-0.58	-0.59	-0.76	-0.90	-0.61	-0.61	-0.28	
c12_Governance_strategies, agendas and plans	-0.38	-0.41	-2.82	-1.51	0.38	-0.58	-0.65	0.53	0.69	3.14	-1.05	0.92	19.2%
c13_Direct financial support for innovation_grants & IPR	0.56	-0.14	0.86	2.61	-0.02	3.14	2.23	1.02	-1.27	-1.76	-3.04	-2.68	18.2%

Source: authors' analysis of STIP Compass (2021 edition), download 19.02.2022

Table 2 STIP compass policy mixes: distribution of standardised chi-squared residuals, by year (2010–2021), and share of policy initiatives (histogram on the right). The black figures highlight significant values of chi-squared residuals (greater or less than 2), whilst the white figures are not significant values; cell background colours: blue scale for positive values; red scale for negative values

groups	no. of terr. ent.s	territorial entities Name	no. pol. Init.s	territorial entities Name	no. pol. Init.s	Characteristic policy mixes labels and test value
a	11	Turkey Portugal Malaysia Slovenia Thailand Ukraine	236 226 166 104 95 74	Kazakhstan Estonia Latvia Bosnia and Herzegovina Slovak Republic	70 54 51 33 26	c4_Human resources_school students c12_Governance_strategies, agendas and plans c6_Finance for start-ups_SMEs
b	23	Germany United Kingdom Brazil France Norway Italy Hungary South Africa Netherlands Finland Colombia Costa Rica	219 198 163 155 151 131 127 119 109 108 104 104	New Zealand Peru Denmark Lithuania Malta Sweden China Argentina Cyprus Czech Republic Iceland	102 92 85 85 84 72 64 59 47 46 31	c12_Governance_strategies, agendas and plans
c	10	United States Australia Canada Austria Poland Spain	219 205 199 174 171 166	Ireland European Union Japan Switzerland	140 125 92 81	c2_Human resources_researchers c10_Research_renewable energy c11_Grants and research funds c3_Human resources_gender equality c8_Ethics
d	13	Russian Federation Korea Israel Chile Belgium - Flanders Luxembourg Greece	151 147 98 94 79 71 66	Belgium - Wallonia Belgium - Brussels Capital Serbia Montenegro Romania Republic of North Macedoı	56 45 41 35 34 17	c13_Direct financial support for innovation_grants & IPR c6_Finance for start-ups_SMEs c7_Policy evaluation
e	1	Croatia	43			c9_Open_science-open_access
f	4	Belgium - Wall.-Bruss Mexico Bulgaria Belgium - Fed. Gov.	42 37 35 32			c9_Open_science-open_access c11_Grants and research funds c1_Indirect financial support for innovation

0 1 2 3 4 5 6 7
Test value

Source: authors' analysis of STIP Compass (2021 edition), download 19.02.2022

Table 3 STIP compass: portfolios of policy mixes on STI policies: cross-territory comparison

STI policy debate before Covid-19, such as research into renewable energy (c10), in 2020 and 2021, ethics (c8), in 2019 and 2020, gender equality (c3), in 2021.

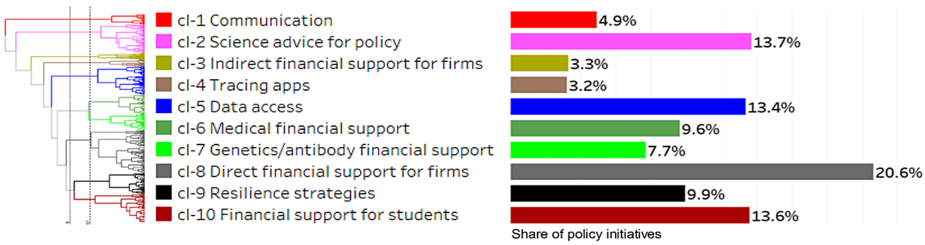
4.1.2 Cross-country comparison based on STIP compass policy portfolios

These thirteen categories of STI policy mixes can be used to characterise STI policy landscapes. With a multidimensional analysis of the matrix ⟨62 territorial entities× *13 clusters of policy mixes*⟩ and a hierarchical cluster analyses of the resulting first ten factors, we obtain the policy portfolios of six groups of territorial entities, listed in Table 3. This table shows the distinctive country policy mixes, identified with their test value (the detailed composition by policy mix and by country is available in Supplementary Table S5, while Supplementary Figure S6 and Figure S7 show, respectively, the factorial map *f1f2* obtained with the correspondence analysis and the dendrogram of the cluster analysis). The focus here is on the similar structural composition of countries, i.e. their STI policy portfolios, defined through the variability in the matrix under analysis with respect to their STI policy mixes. The six groups of policy portfolios have a share of policy initiatives almost proportional to the number of countries in the group and do not belong to specific geographical areas. The largest group of countries (group-b, 37% of countries and 39% of policy initiatives) is essentially characterised by the barycentre STI policy mix: the one oriented by the national government (c12). Group-a, group-c and group-d have almost the same number of countries (respectively, 11, 10 and 13). In the case of group-a, the share of countries is almost proportional to the share of policy initiatives, with distinctive policy mixes embracing governance (c12), but mainly support for students (c4) and, to a lesser extent, finance for start-ups (c6). Group-c and group-d have opposite relative shares of number of countries and policy initiatives: the biggest share of policy initiatives, 25%, distinctively characterises countries in group-c, mainly targeting researchers (human resources_researchers, c2), a focus on renewable energy (c10), funding of research through grants (c11) and a specific initiative to reduce gender inequalities in research (c3), by controlling for ethics in innovation processes (c8), that we have seen is an issue specifically related to AI. Countries in group-d have a relatively lower share of policy initiatives (15%), with a more traditional innovation policy focus: finance (direct financial support for innovation grants and IPR, c13, and finance for start-ups, c6) and policy evaluation (c7). The two smallest groups group-e and group-f have a common focus on open innovation (distinctively characterising Croatia's policies, in group-e), while the four territorial entities in group-f have two other areas of distinctive, quite traditional, policy interventions: grants and research funds (c11) and indirect financial support for innovation (c1).

4.2 STIP covid-19 watch

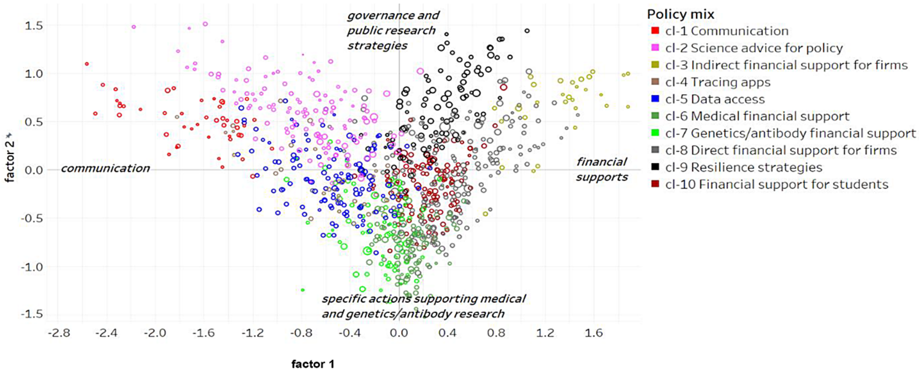
4.2.1 STI policy mix

In this group of STI policy initiatives, the optimal number of clusters would be eight (see the Caliński-Harabasz index, Supplementary Figure S8), but the ten cluster cut (see Fig. 5) allows us to split a larger group of interventions in the medical area. The graphical results of clustering on codes and terms are shown in the factorial map *f1f2* in Fig. 6: analogously to Fig. 4, each dot in the factorial map is a policy initiative document, coloured according to



Source: authors' analysis of STIP Covid-19 Watch, download 19.02.2022

Fig. 5 STIP Covid-19 Watch: Dendrogram of the hierarchical cluster of policy initiatives, classification based on multiple coding and terms (on the left), legend of colours and labels of the clusters (center), and share of policy initiatives by cluster (histogram on the right). Share of policy initiatives. Source: authors' analysis of STIP Covid-19 Watch, download 19.02.2022



Source: authors' analysis of STIP Covid-19 Watch, download 19.02.2022

Fig. 6 STIP Covid-19 Watch: Factorial map *f1f2* - Distribution of Policy Initiatives - Matrix (*935 Policy Initiatives* × *126 Codes&827 selected Terms*) (935×953). Each dot corresponds to a policy initiative, with size proportional to the number of codes and terms. Dot colour is by policy mix detected by clustering. factor 1. Source: authors' analysis of STIP Covid-19 Watch, download 19.02.2022

the policy mix cluster in which the document has been classified by the clustering algorithm, and the proximity of dots represents the similarity with respect to the various codes and terms used in the STIP Covid-19 Watch dataset. Cluster labels were assigned by reading the characteristic codes and terms in each cluster (listed in Supplementary Table S6 in decreasing order of their test value). As an example of labelling for this clustering, let us consider cl-9, "resilience strategies" (quotations from characteristic terms and codes are in italics):

> The distinctive terms of this group of innovation policies embrace the many dimensions of resilience in economies and societies. In decreasing order of test value (from Supplementary Table S6), the characteristic terms are *investment*, *transition*, *digital*, *green*, *sustainable*, *recovery*, *economy*, *reform*, *plan*, *structural*, *resilient*, *women*, *average*, *participation*, *strategic*, *education*, *growth*, *and competitiveness*.
>
> These policies are characterised, first of all, by a set of related themes: *building more resilient societies and economies* (TH98) and an *STI plan or strategy* (TH13), countering the impacts of Covid-19 on STI systems through *mitigating the long-term impacts of Covid-19* (TH97), a *research and innovation for society strategy* (TH58), *public research strategies* (TH18), *business innovation policy strategies* (TH30), *European Research Area (ERA)-related strategies* (TH70).
>
> Mainly, there is just one policy instrument: governance through *strategies*, *agendas and plans* (PI024). The target groups embrace the *labour force in general* (TG19); intermediaries, in particular, *industry associations* (TG36), *technology transfer offices* (TG35), *incubators*, *accelerators*, *science parks or techno parks* (TG34); social groups are especially emphasised and in particular *women* (TG14) and *disadvantaged and excluded groups* (TG15), and *private research and development labs* (TG22).

The cloud of dots in Fig. 6 highlights a marked polarisation of policy mixes along the two factors (Supplementary Figure S9 shows the combination of pairs for the first ten factors to interpret the polarisation). Along the first factor (from left to right), the polarisation unfolds from policy measures to provide knowledge and communication, vs. funding and financial support. Policies targeting communication, science advice for policy, and data access (respectively, cl-1, cl-2 and cl-5), placed on the left side of the first factor, are opposite to the policies targeting direct and indirect financial support for firms (respectively, cl-8 and cl-3), on the right of the first factor.

Along with the second factor, the polarisation (from bottom to top) is between specific actions supporting medical and genetics/antibody research vs. governance and public research strategies. The first group of policy mixes (on the bottom) focuses on policies that target medical (cl-6) and genetics/antibody (cl-7) financial support while the second group focuses on science advice for policy and resilience strategies (respectively, cl-2 and cl-9).

4.2.2 Cross-country comparison based on STIP Covid-19 watch policy portfolios

The overall STI policy landscape emerging from the identification of policy mixes in the Covid-19 specific policies can be discussed concerning a country's respective STI policy portfolios. Through the multidimensional and cluster analysis on the matrix ⟨40 *territorial entities*× *10 clusters of policy mixes*⟩[4], we identify seven groups of countries.

Analogously to the STIP Compass, the similar structural composition of countries concerning their STI policy mixes is interpreted here in terms of the characteristic policy mixes weighted by their test values (Table 4), while the detailed composition by policy mix and by country is available in Supplementary Table S7. Supplementary Figures S10 and S11 show, respectively, the factorial map *f1f2* obtained with the correspondence analysis and the dendrogram of the cluster analysis.

The STI policy portfolio of the largest group of 11 countries (group A) has implemented policies focusing both on direct financial support and on advice for science. The policy portfolio of the second largest group of nine countries (group D) combines resilience strategies and policies to supply financial support to students. The distinctive policy portfolios of the group of six countries plus the European Union (group B) is characterised by knowledge and communication and data access policies together with resilience strategies. A group of four countries (group C) largely relies on financial support (both indirect and direct); while another group of four countries (group E) focuses policies on genetics/antibody financial support and data access. In a group of three countries (group F), medical financial support prevails over other policy mixes; while two countries (group G) are mainly implementing policies on data access, and also on tracing apps and on indirect financial support.

Except for group B, every territorial entity places reasonable emphasis on financial support, either with a specific target domain (group E) or with direct financial support (groups A, C, D, F) or indirect financial support (group G). Group B features a stronger focus on communication and science advice however it should be noted that this group includes the

[4] The following territorial entities with fewer than 10 policy initiatives have not been included in the analysis because the number of policies is too small: Israel (8 policy initiatives), Colombia, and Thailand (6 policy initiatives), Iceland and Mexico (5 policy initiatives), Belgium - Brussels Capital, Belgium - Wallonia-Brussels Federation, Cyprus and Malta (4 policy initiatives), Belgium - Wallonia, Bosnia and Herzegovina and Romania (3 policy initiatives), Hungary, India, Kazakhstan, Serbia and Ukraine (2 policy initiatives), Belgium - Flanders, Croatia, Luxembourg, Montenegro (1 policy initiative).

groups	n. of terr.e nt.s	territorial entities Name	n. pol. Init.s	territorial entities Name	n. pol. Init.s	Characteristic policy mixes labesl and test-value
A	11	Canada United States Finland Switzerland Belgium	81 42 32 24 19	Australia Chile Czech Republic Estonia Peru New Zealand	17 17 17 14 14 13	cl-8 Direct financial support for firms cl-2 Science advice for policy cl-4 Tracing apps
B	7	European Union Portugal Germany Korea Spain	44 35 31 29 21	South Africa Denmark	16 12	cl-1 Communication cl-9 Resilience strategies cl-5 Data access cl-2 Science advice for policy
C	4	Norway Russian Federation Italy Netherlands	32 30 16 13			cl-3 Indirect financial support for firms cl-8 Direct financial support for firms cl-1 Communication cl-9 Resilience strategies
D	9	Turkey Lithuania Malaysia Argentina Brazil Sweden	29 22 18 14 14 14	Costa Rica Ireland Belgium-Federal gov.t	11 11 10	cl-9 Resilience strategies cl-10 Financial support for students
E	4	Greece Slovenia Japan Austria	26 17 14 13			cl-7 Genetics/antibody financial support cl-5 Data access cl-9 Resilience strategies
F	3	Poland United Kingdom France	21 19 17			cl-6 Medical financial support cl-7 Genetics/antibody financial support
G	2	China Latvia	16 11			cl-5 Data access cl-4 Tracing apps cl-3 Indirect financial support for firms

0 5 10 15 20 25 30 35
Test value

Source: authors' analysis of STIP Covid-19 Watch, download 19.02.2022

Table 4 STIP Covid-19 Watch: portfolios of policy mixes on STI policies: cross-territory comparison. Territorial entities with at least 10 policy initiatives

European Union and Denmark, Germany, Portugal and Spain which are EU members and thus involved in broader discussions inside the EU, at least for some policy coordination among member states in addition to measures taken by the European Commission. Thus we cannot exclude an overlap or a duplication of policy measures which are based on EU-wide measures and converted into domestic measures.

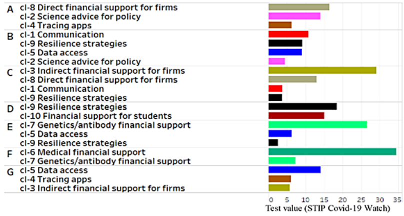
4.3 Cross-country comparison on the policy mixes of the STIP Compass and STIP Covid-19 Watch

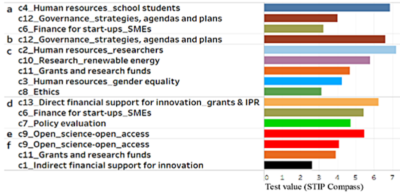
A comparison of the different STI policy portfolios of the territorial entities, obtained through the analyses of the STIP Compass and the STIP Covid-19 Watch (Table 5), allows research question no.3 to be answered. Although somewhat cumbersome, the cross-country comparison of the STI policy portfolios presented in a contingency table returns a more significant result than the one that might be computed with a cluster analysis of the countries classified according to the two sets of policy portfolios: only 39 countries would enter into such an analysis that would refer to 23 portfolios (13 for non-Covid-19 related policies and 10 for the Covid-19 specific STI policies).

Apart from the countries that cannot be compared because they appeared in the STIP Covid-19 Watch dataset with fewer than 10 policy initiatives, the results show a variety of policy portfolios implemented to cope with the Covid-19 pandemic, by groups of countries that had similar policy portfolios in the non-Covid-19 related STI policies.

The largest group of countries (Group b) that had a STIP Compass policy portfolio essentially characterised by the policy mix governance_strategies, agendas and plans (c12) addressed the pandemic with an array of specific policies, that cannot be explained by the relative severity of the pandemic in the country. When measured by the excess mortality

Groups STIP Covid-19 Watch	Groups STIP Compass a	b	c	d
A	EST	CZE FIN NZL PER	AUS CAN CHE USA	CHL
B	PRT	DEU DNK ZAF	ESP EU	KOR
C		ITA NLD NOR		RUS
D	MYS TUR	ARG BRA CRI LTU SWE	IRL	BEBRU
E	SVN		AUT JPN	GRC
F		FRA GBR	POL	
G	LVA	CHN		





Source: authors' analysis of STIP Compass (2021 edition) and STIP Covid-19 Watch, download 19.02.2022

Table 5 Cross-territory comparison of Covid-19 specific (STIP Covid-19 Watch) and non-covid-19 related (STIP compass) STI innovation policy portfolios. Territorial entities with at least 10 policy initiatives

rate, we observe that countries with similar impacts of the pandemic (see OECD 2021; 2023) had very different STI policy interventions. An explanation of the differences in Covid-19 specific STI policies has to be explained by countries' structural weaknesses and strengths, the vocation/competence networks of the country, and also the consolidated policy design in the country (such as the use of financial policy instruments). Considering the case of four countries in such groups, France, Germany, Italy, and the United Kingdom, which have comparable economies and very high numbers of infected individuals, Covid-19 specific STI policies in France and the United Kingdom are characterised by medical financial support (with the UK in the frontline in developing a vaccine, leveraging EU funds and the outstanding research infrastructure in medical research), while Germany was adopting policies characterised by knowledge and communication and data access policies together with resilience strategies, and Italy is in the group of countries that distinctively relies on indirect and direct financial support.

Another example can be taken with the group of four countries Austria, Greece, Japan, and Slovenia that share a similar landscape of Covid-19 specific policy initiatives, focusing mainly on medical research and data access. Socio-economic conditions in those countries, however, are very different, as are their longer-term STI policy portfolios. To explain the specific focus of Greece, it is worth noting its traditional policy portfolios but also a longstanding focus on medical research that becomes specific in the pandemic, while Austria and Japan have, in their long track record of STI policies, a wide range of STI policies, which they relied on during the pandemic, complementing them with policies to support pioneering medical research.

5 Limitations and strengths of the analysis

The data and methods proposed in this paper have some limitations but also strengths that should be kept in mind in the concluding remarks and the discussion of implications for policy design and analysis.

Concerning the data source, the completeness and consistency of the data provided by the OECD STIP Lab have been improved over previous editions of STIP Compass, but some relevant information is still missing, as in the case of the budget range, which is available for 70% of the policy initiatives, and this makes the comparison of policies weaker. In addition, although the classification of innovation policy dimensions provided by the OECD is robust, in the cross-country comparison, we are assuming that the National Contact Points share the same interpretation of that classification and that when they enter the information in the STIP Compass dataset they share the same degree of accuracy in referring to the actual range of innovation policy initiatives implemented in their country. Also, the information included in the datasets does not allow the STI policy measures to be connected to the actual STI landscape (infrastructure) of the respective country. A policy measure's effectiveness and efficiency are likely to depend on the respective national institutional landscape. This however, cannot be analysed with the given dataset.

Concerning data, a second issue is the mix of territorial entities: the country level, the sub-national scale (in the case of Belgium), and the supranational scale, in the case of the European Union. Notwithstanding its contribution to broadening the variety of information, this mix, on one side, does not make it possible to compare territorial entities at the sub-national level, because the data refers only to Belgium, and, on the other side, the only supranational entity, the European Union, should not be treated as a country. For these reasons, sub-national entities have been omitted from the cross-country comparison, while the EU has been kept with the aforementioned caveat. As already stressed by Russo and Pavone (2021), overcoming the limited information on sub-national STI policies is possible, at least for EU countries, that ordinarily collect information on respective policies at the regional level (as they receive EU funds and must be applied to the EU for final approval of the resources). In this case, it is the European Commission that should orient the sub-national collection of information more concretely.

The two sets of policies examined in this paper - STIP Compass and STIP Covid-19 Watch - have a sharp characterisation because of the specific lexicon and ad hoc policy themes adopted in the Covid-19 specific policy initiatives, but also in terms of policy instruments and target beneficiaries (Fig. 1). The STIP Covid-19 Watch policy initiatives are particularly notable for their characteristic policy instruments that specifically leverage the role attributed to science in coping with the pandemic. These initiatives have a broad range of beneficiaries, including society, public and private sectors, demonstrating their inclusivity. STIP Compass, as observed by Russo and Pavone (2021), shows a variety of other dimensions. However, during the pandemic (2020 and 2021), it shares similarities with the Covid-19 Watch STI policy initiatives, such as a relatively higher emphasis on policy instruments focusing on science and technology regulation, the involvement of governmental entities at international, national and sub-national levels, and a focus on relatively more vulnerable beneficiaries in the pandemic, such as firms of all sizes and women.

A major strength of the analysis is the focus on each dataset of policy initiatives (respectively, STIP Compass and STIP Covid-19 Watch) to pick out the specificities of Covid-19

STI policies: this choice makes it possible to identify 10 policy mixes, while the aggregation of the two datasets singles out only two. This result allows a richer understanding of the Covid-19 specific STI policies in the two years of the emergency phase of the pandemic (2020 and 2021).

Concerning a previous analysis on STIP Compass (Russo and Pavone 2021; Howoldt 2024), in this paper we have improved the readability of semantic analysis by applying the multidimensional analysis to codes and terms, thus allowing finer-grained and more consistent groups of elements characterising the various policies mixes to be identified. The online tool[5] allows easy access to the analysis's results by focusing on each set of STI policies and their codes and terms.

6 Concluding remarks

The large number of elements that characterise STI policies in terms of themes, instruments, target groups of beneficiaries (altogether, about 100 codes and more than three thousand terms used to describe the policies) gives an account of a great variety of aspects concerning policy actions in support of science, technology and innovation. Through a semantic analysis of the textual content and codes of documents describing STI policies, in this paper, we have built a dataset and analysis tools to compare countries regarding their non-Covid-19 related and Covid-19 specific policy portfolios. The results allow, first of all, the identification of policy mixes in the two sets of policy initiatives, defined with a vector space model and a hierarchical cluster analysis, as statistically significant combinations of terms and codes: 13 policy mixes of non-Covid-19 related STI policies (with different characterisation by year, as seen in Table 2) and 10 policy mixes of Covid-19 specific ones (which characterise the STI policies implemented by countries to cope with the pandemic). The results are the answer to the first research question.

We then answered the second research question with the identification of cross-country similarities in policy portfolios. In particular, we have pointed out that, in the OECD countries, there are six groups of countries with similar policy portfolios, in the case of non-Covid-19 related STI policies, and seven groups of countries with similar policy portfolios, in the Covid-19 specific STI polices. Each group is identified in terms of a characteristic policy mix. In the case of non-Covid-19 related policies (Table 3), the largest group of countries is essentially characterised by the barycentre policy mix, with policies oriented by the national government, and the other five groups of countries show combinations ranging from a more traditional focus on governance, support for students, finance for start-ups to targets specific to researchers, technological domains such as renewable energy, specific initiatives to reduce gender inequalities in research, by controlling for ethics in innovation processes, specifically concerning AI. When considering Covid-19 specific STI policies (Table 4), most countries have, mainly, a combination of policies over the three domains of financial support, government strategies, and knowledge and communication, with only minor policies initiatives on the medical domain; one-fourth of the countries combines medical & genetics with knowledge and communication; while only three countries focus their policy mix portfolio on medical and genetics themes.

[5] https://public.tableau.com/views/CrossCountryAnalysis/STIPolicies?:language=en-US&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link.

The cross-country comparison on the similarities between non-Covid-19 related and Covid-19 specific STI policy portfolios (the third research question) highlights groups of countries with similar structural compositions of their innovation policy portfolios and a range of strategies in addressing the pandemic with STI policy initiatives (Table 5). This cannot be explained only by the relative severity of the pandemic in the country. One should address the specific structural weaknesses and strengths, the vocation/competence networks of the country, but also the consolidated policy design in the country (such as the use of financial policy instruments).

The debate on innovation policies during the pandemic has stressed the areas of interventions: medical research, data access and open data, financial support, and government coordination of interventions. The evidence-based analysis of STI policy portfolios proposed in this paper provides an original contribution to defining which policy mixes characterise the STI policies implemented by OECD countries to cope with the pandemic and the corresponding classification of countries' innovation policy initiatives. The ten groups of identified policy mixes cover four main areas:

• The largest group of policy initiatives, almost 37%, specifically covers direct and indirect financial support for firms (mainly SMEs and young firms) and students:

- A second group, 23.6% of the policy initiatives, concerns governance and public research strategies (characterised by resilience strategies and science advice for policy).
- A third group, 21.5% of the policy initiatives, focus on knowledge and communication aspects, such as communication, data access and tracing apps
- The fourth group, 17.3% of the policy initiatives, targets the specific actions supporting medical and genetics/antibody research.

The first two groups appear to show similar features to policy response to any form of crisis. Similar responses were observed during the financial crisis of 2008 and other smaller crises. What is new this time is the emphasis on communication strategies and governance, targeted at science advice and the establishment of respective councils or policy advisory bodies. From today's perspective, it is impossible to assess the impact of scientific advice and the establishment of councils that were set up in the emergency phase. Indeed, the public became somewhat confused when listening to various scientists and hearing different thoughts and opinions on the same issue. Typically, scientists do have different opinions and perceptions on any one issue, but obviously the public is not always aware of such reasonable discussions within a community. Politicians, on the other hand, appointed numerous councils and advisory bodies to legitimise their decisions and inform the public about these decisions which they submitted as being science-based and sometimes even evidence-based. The scientific community found itself in the uncomfortable position of being in the middle of two sometimes diverging demands, with pressure from politicians delivering evidence to justify measures and the public, which put pressure on scientists to use their influence on politicians regarding unpopular measures, such as lockdown or vaccination mandates. Both sides neglected the fact that there was no evidence base available. The shift in funding priorities towards medical and genetics/antibody research comes with little surprise, especially in light of the UK and France, which possess a reasonable pharmaceutical industrial base. The question, however, is whether countries invested additional funding in this field or whether they shifted existing funds from other fields/topics of research. Even though the information available in the datasets does not allow this analysis, we observed that, during the years 2020 and 2021, the non-Covid-19 related STI policies (classified in the STIP Compass) were similar to the ones implemented in the previous years (Fig. 1, panel c). The additional funding can be inferred by the implementation of Covid-19 specific STI policies.

7 Discussion

In the scientific community, it is largely accepted that in a cross-country comparison the multidimensional nature of STI policies cannot be considered without a clear reference to the complex weaving of public sector structures and the cumulative amount of resources for STI policies over time, including socio-economic and technological conditions.

Now that the pandemic is over, having identified which portfolios of policy mix characterised countries in terms of their non-Covid-19 related and Covid-19 specific STI policies fills a knowledge gap on changes in STI policies induced by exogenous shocks.

Beyond being aware of the difficulties in such an analysis, there is a minimum goal that can be achieved by policy makers: they can look at their Covid-19 innovation policies through the lenses of the multidimensional categories of policy mixes analysed in this paper. A validation of the results would also open up a further step in a common understanding of the use of new tools to address multidimensional phenomena (such as, for example, the statistical tools applied to semantic analysis and to clustering proposed in this paper, or the methods implemented by Russo and Pavone (2021) or by Howoldt (2024), on STIP Compass data). Framing policy measures based on the lenses of multidimensional and multilevel interacting elements is a challenge that policy makers should consider when they design, implement and assess the impact of the policies. And this is still an open challenge, not only concerning innovation policies.

When we move to the comparison of non-Covid-19 related and Covid-19 specific STI policy portfolios, we do not find a clear cross-country pattern. In general, although a variety of responses to the Covid-19 pandemic is a result, the cross-country comparison is still open to the identification of paths of analysis to interpret why a country's policy landscape has changed – and, in different ways – among countries with similar pre-Covid policy landscapes.

The cross-country comparison emerging from the analysis proposed in this paper can be further explored with complementary information on the assessment of those policies and on the structural socio-economic features characterising the countries with respect to the relevant STI aspects under analysis.

Taken as a contingent need to cope with the current pandemic, Covid-19 specific STI policies will contribute to the design and implementation of policies supporting the transition toward all-pandemic resilient societies, an issue that is at the core of the Next Generation EU as well as of the post-pandemic recovery plans in most countries, all over the world. The OECD STIP Compass will continue to support data for such a comparison and we suggest that the implementation of multidimensional methods of data analysis should be added to the OECD toolkit, to be used in describing the ongoing changes occurring in the innovation policy domain. The application to further editions of the policy documents in the STIP Compass of the same multidimensional methods of analysis will provide insights into the changing policy mixes occurring over time.

Acknowledgements Authors wish to thank OECD, for making available the complete download of STIP Covid-19 Watch dataset, Caroline Paunov and the participants to presentation of the paper at the *fteval* European R&I Policy Evaluation Conference (5–6 May 2022, Vienna, Austria) for their comments to a preliminary version of this paper.

Author contributions The contribution of the authors in writing the paper is as follows: Margherita Russo and Dirk Meissner: Conceptualization, Project administration, Resources, Supervision. Pasquale Pavone and Fabrizio Alboni: Data Curation, Formal analysis, Methodology, Software. Margherita Russo, Pasquale Pavone and Fabrizio Alboni: Investigation, Validation, Visualization, Writing - Original Draft, Writing - Review & Editing.

Funding Open access funding provided by Università degli Studi di Modena e Reggio Emilia within the CRUI-CARE Agreement. The paper is a contribution of the Project "*Textpol*. Textual analysis tools for the design, analysis, and monitoring of public policy", FAR 2021, Dipartimento di Economia Marco Biagi, Università di Modena e Reggio Emilia, Italy. Dirk Meissner's contribution to this article is based on the study funded by the Basic Research Program of the HSE University.

Data availability The datasets analysed in this study are available in the STIP Compass edition (EC-OECD 2023a), the STIP Covid-19 Watch (EC-OECD 2023b), download 19.02.2022.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

References

Abi Younes, G., Ayoubi, C., Ballester, O., Cristelli, G., de Rassenfosse, G., Foray, D., Gaulé, P., et al.: COVID-19: Insights from Innovation Economists. Sci. Public. Policy. **47**(5), 733–745 (2021). https://doi.org/10.1093/scipol/scaa028

Alboni, F., Pavone, P., Russo, M.: The search for topics related to electric mobility: A comparative analysis of some of the most widely used methods in the literature. METRON, 1–25. (2023)

Antons, D., Grünwald, E., Cichy, P., and Salge T. O.: The application of text mining methods in Innovation Research: Current state, evolution patterns, and Development priorities. R&D Manage. **50**(3), 329–351 (2020). https://doi.org/10.1111/radm.12408

Baute, S., Anniek de Ruijter: EU Health Solidarity in Times of Crisis: Explaining Public preferences towards EU Risk Pooling for Medicines. J. Eur. Public. Policy. **0**(0), 1–23 (2021). https://doi.org/10.1080/13501763.2021.1936129

Boin, A. and Lodge, M.: Responding to the COVID-19 Crisis: A principled or Pragmatist Approach? J. Eur. Public. Policy. **28**(8), 1131–1152 (2021)

Bolasco, S., Baiocchi, F., Canzonetti, A., De Gasperis, G.: TaLTaC3. 0, Un software multi-lessicale e unitestuale ad architettura web. In Proceedings of 13th International Conference on Statistical Analysis of Textual Data, Eds D. Mayaffre, C. Poudat, L. Vanni, V. Magri, and P. Follette Nice: University Nice Sophia Antipolis, 7–10 (2016)

Bozeman, B.: Use of Science in Public Policy: Lessons from the COVID-19 pandemic efforts to 'Follow the science'. Sci. Public. Policy. **49**(5), 806–817 (2022). https://doi.org/10.1093/scipol/scac026

Caliński, T. and Harabasz J.: A Dendrite Method for Cluster Analysis. Commun. Statistics-Theory Methods. **3**(1), 1–27 (1974)

Capano, G. and Lippi A.: Decentralization, policy capacities, and varieties of First Health response to the COVID-19 outbreak: Evidence from three regions in Italy. J. Eur. Public. Policy. **0**(0), 1–22 (2021). https://doi.org/10.1080/13501763.2021.1942156
Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A.: NbClust: An R Package for determining the relevant number of clusters in a Data Set. J. Stat. Softw. **61**, 1–36 (2014)
Dziembała, M.: Science, Technology and Innovation Policy in the USA During the Covid-19 Pandemic. In *Research and Innovation Forum 2022*, edited by Anna Visvizi, Orlando Troisi, and Mara Grimaldi, 83–95. Cham: Springer International Publishing. (2023). https://doi.org/10.1007/978-3-031-19560-0_6
EC-OECD. STIP Compass Taxonomies Describing STI Policy Data. (2021). https://stip.oecd.org
EC-OECD: STIP COVID-19 Watch | STIP Compass. 2023. (2023b). https://stip.oecd.org/covid/
EC-OECD: Download Data | STIP Compass. 2023. (2023a). https://stip.oecd.org/stip/download-data
Gershman, M., Gokhberg, L. and Kuznetsova T.: Science, Technology, and Innovation Policy in Response to the COVID-19 Crisis: The Case of Russia. In *COVID-19 and Society: Socio-Economic Perspectives on the Impact, Implications, and Challenges*, edited by Mustafa Polat, Serhat Burmaoglu, and Ozcan Saritas, 279–300. Cham: Springer International Publishing. (2022). https://doi.org/10.1007/978-3-031-13142-4_14
Gokhberg, L., Meissner, D. and Kuzminov I.: What semantic analysis can tell us about Long Term trends in the global STI policy agenda. J. Technol. Transf., 1–29. (2022)
Howoldt, D.: Characterising Innovation Policy mixes in Innovation systems. Res. Policy. **53**(2), 104902 (2024)
Lê, S., Josse, J., François, Husson: FactoMineR: An R Package for Multivariate Analysis. J. Stat. Softw. **25**, 1–18 (2008)
Lebart, L., Salem, A., Berry, L.: Exploring Textual Data. Springer, Dordrecht; London (1998)
Meissner, D. and Kergroach S.: Innovation Policy Mix: Mapping and Measurement. J. Technol. Transf. **46**(1), 197–222 (2021). https://doi.org/10.1007/s10961-019-09767-4
Narlikar, A. and Sottilotta, C. E.: Pandemic narratives and policy responses: West European governments and COVID-19. J. Eur. Public. Policy. **0**(0), 1–20 (2021). https://doi.org/10.1080/13501763.2021.1942152
OECD: Ready for the Next Crisis? Investing in Health System Resilience. 2023. (2023). https://www.oecd.org/health/ready-for-the-next-crisis-investing-in-health-system-resilience-1e53cf80-en.htm
OECD: Excess Mortality since January 2020. November 17, 2021. (2021). https://www.oecd.org/coronavirus/en/data-insights/excess-mortality-since-january-2020
OECD: STI Policy and the COVID-19 Crisis - OECD. 2020. (2020). https://www.oecd.org/sti/science-technology-innovation-outlook/STIP-and-COVID-19-crisis/
OECD: STIP Data Lab | STIP Compass. Accessed May 2, 2024. (2024). https://stip.oecd.org/stip/pages/stipDataLab
Pavone, P.: Automatic Multiword Identification in a Specialist Corpus. In *Tracing the Life Cycle of Ideas in the Humanities and Social Sciences*, edited by Arjuna Tuzzi, 151–66. Cham: Springer International Publishing. (2018). https://doi.org/10.1007/978-3-319-97064-6_8
Plümper, T. and Neumayer E.: Lockdown policies and the dynamics of the First Wave of the Sars-CoV-2 pandemic in Europe. J. Eur. Public. Policy. **29**(3), 321–341 (2022)
Russo, M. and Pavone P.: Evidence-based portfolios of innovation policy mixes: A cross-country analysis. Technol. Forecast. Soc. Chang. **168**, 120708 (2021)
Salton, G., Wong, A., Chung-Shu, Y.: A Vector Space Model for Automatic indexing. Commun. ACM. **18**(11), 613–620 (1975)
Schmidt, V.A.: European Emergency Politics and the question of legitimacy. J. Eur. Public. Policy. **0**(0), 1–15 (2021). https://doi.org/10.1080/13501763.2021.1916061
Toshkov, D., Carroll, B. and Yesilkagit K.: Government Capacity, Societal Trust or Party preferences: What accounts for the Variety of National Policy responses to the COVID-19 pandemic in Europe? J. Eur. Public. Policy. **0**(0), 1–20 (2021). https://doi.org/10.1080/13501763.2021.1928270
Ward, J.H.: Hierarchical grouping to optimize an objective function. J. Am. Stat. Assoc. **58**(301), 236–244 (1963). https://doi.org/10.1080/01621459.1963.10500845
WHO, World Health Organization: Coronavirus Disease (COVID-19) Pandemic. World Health Organization. 2024. (2024). https://www.who.int/europe/emergencies/situations/covid-19
Yun, T.: Review of Science and Technology Innovation policies in Major innovative-oriented countries in response to the COVID-19 pandemic. Biosaf. Health. **5**(01), 8–13 (2023)

Authors and Affiliations

Margherita Russo[1] · **Pasquale Pavone**[2] · **Dirk Meissner**[3] · **Fabrizio Alboni**[4]

✉ Pasquale Pavone
pasquale.pavone@unimore.it

Margherita Russo
margherita.russo@unimore.it

Dirk Meissner
dmeissner@hse.ru

Fabrizio Alboni
fabrizio.alboni@unibo.it

1 Dipartimento di Economia Marco Biagi and CAPP - Research Centre for the Analysis of Public Policies, Università di Modena e Reggio Emilia, Modena, Italy

2 CAPP - Research Centre for the Analysis of Public Policies, Università di Modena e Reggio Emilia, Modena, Italy

3 National Research University Higher School of Economics, Russian Federation, Moscow, Russia

4 Dipartimento di Statistica, Università di Bologna, Bologna, Italy