The study of policy coordination: an approach to integrate expert assessment with automated content analysis

Peer-reviewed author version

KELCHTERMANS, Stijn & Santos, Anabela M. (2024) The study of policy coordination: an approach to integrate expert assessment with automated content analysis. In: Peace economics, peace science, and public policy, 52 (2), p. 222-235.

DOI: 10.1093/scipol/scae074
Handle: http://hdl.handle.net/1942/44889

The Study of Policy Coordination:
An Approach to Integrate Expert Assessment with Automated Content Analysis

Stijn Kelchtermans[*,**] (corresponding author)

Anabela M. Santos[***,****]

Abstract

Public policy scholars have long agreed that the coordination of policies involving different authorities and policy frameworks is challenging. This paper responds to recent calls to explore new empirical methods for studying policy coordination by integrating automated text analysis into the workflow of human experts. A key contribution of our approach is a quantitative indicator of the semantic alignment between different policy objectives. We critically review the performance of the indicator in terms of its ability to assess both horizontal and vertical policy coordination by analysing the Recovery and Resilience Plan (2021-2026) and the national and regional Smart Specialisation Strategies (2021-2027) in Portugal. Based on this exercise, we reflect on the necessary conditions for the successful implementation of automated text analysis of policy documents as well as its potential in other empirical settings.

Keywords: Policy Coordination; Text Analysis; Research and Innovation; Smart Specialisation; Recovery and Resilience Plan

[*] Department of Management, Strategy & Innovation, KU Leuven, Warmoesberg 26, 1000 Brussels, Belgium, +32 2 608 14 39 stijn.kelchtermans@kuleuven.be

[**] Department of Marketing & Strategy, Hasselt University, Belgium

[***] Joint Research Centre, European Commission, unit B3, Seville, Spain, Anabela.MARQUES-SANTOS@ec.europa.eu

[****] Disclaimer: The views expressed are purely those of the author and may not in any circumstances be regarded as stating an official position of the European Commission.

Funding: the work of Stijn Kelchtermans was supported by the KU Leuven Impulsfund (grant number IMP/20/002).

1. Introduction

Policy coordination refers to the design of policy measures by different government entities such that they are consistent and synergetic, or at least do not conflict with each other (Boston, 1992). Although desirable from the point of view of efficiency and quality of public service, there is a long-standing consensus among public policy scholars that coordinating policies involving different authorities and policy frameworks is challenging (Cejudo & Michel, 2022). Lack of coordination among policies may arise for multiple reasons, such as the deep specialization of government agencies, or the dispersion of authority among different political parties, including "turf wars" that inhibit the exchange of information (Peters, 2018). Existing research has approached the subject of policy coordination from various angles (Tosun & Lang, 2017), with particular attention paid to country-specific cases that take a political perspective (e.g. Griessen & Braun, 2008; Tamtik, 2017). Previous work has relied strongly on qualitative methods (Trein et al., 2019), while comparatively little attention has been devoted to developing more generally applicable methods for analysing policy coordination, which is the aim of this paper.

More specifically, we address the need for a better assessment of the degree of relatedness of different policy measures. We do not directly evaluate the complementary effects of different policies on a certain outcome, but rather propose an approach for evaluating the extent to which the potential for such synergies is present. In doing so, we respond to recent calls for more research on policy coordination and integration that exploits the possibilities offered by computational text analysis (Trein *et al.*, 2020). Such techniques have recently been applied in diverse domains, like consumer research (Humphreys & Wang, 2018), corporate disclosure statements (Chakraborty & Bhattacharjee, 2020) or firm innovation strategies (Belderbos *et al.*, 2017). However, the use of computational text analysis to analyse government policies and their coordination has so far been limited with the exception of keyword-based approaches, such as Huang et al. (2023) who study themes of policy coordination and how they change over time. A first objective of this paper is to move beyond simple keyword matching by assessing the *semantic similarity* between *pairs of policy documents*, providing the basis for assessing the complementarity between the respective policies. The second objective addresses the caution expressed by policy scholars about the potential pitfalls of automated text analysis despite its benefits, such as the ability to rapidly analyse large volumes of data. Grimmer & Stewart (2013) advocate the use of automated text analysis but at the same time warn that *"such quantitative methods augment humans, not replace them"*. Rather than relying on automated text analysis alone, the method we develop shows how expert assessment and automated text analysis strengthen each other in an integrated approach.

Empirically, we focus on the case study of Portugal and analyse the alignment between the measures of its Recovery and Resilience Plan (RRP), as part of *Next Generation EU* (the temporary instrument to support Covid-19 recovery) and the innovation priorities of Smart Specialisation Strategies (S3).[1] There are two reasons for the selection of the RRP and S3 policies. First, analysing them jointly ties in with the long-standing interest in the European policy realm in monitoring and analysing the overall effectiveness of national and regional R&I policies with the aim of an improved implementation of EU policy goals. In

[1] We provide a brief background on these policies in section 2.

line with that overarching goal, this study contributes to the knowledge of potential synergies between Member States' Smart Specialisation Strategies and their Recovery and Resilience Plans in order to inform the European Semester, the EU's annual cycle of economic policy coordination. In addition to the EU's 2021-2027 Multiannual Financial Framework of €1,210.9 billion, EU Member states also have access to the funds of the *Next Generation EU* instrument, amounting to €806.9 billion (European Commission, 2021a). Alignment of actions in the Recovery Plans with the national and regional policy priorities is considered essential to ensure a higher effectiveness of these public funds. A second reason for selecting the RRP and S3 policies lies in their respective rationales, which create expectations of some degree of alignment. More specifically, the RRP's not only have the short-term objective of reactivating demand, which slowed dramatically due to the Covid-19 crisis, but also contain the long-term goal of structural economic transformation through strategic investments and reforms (European Commission, 2020). Smart Specialisation Strategies put forward medium- and long-term targets of economic transformation and growth, following their signature approach of prioritising investments in selected innovation areas based on regional strengths (Foray *et al.*, 2012). In other words, there is potential for synergies between RRP and S3 policies since investments put forward in the RRP may support the implementation of Smart Specialisation by reducing obstacles to innovation, by targeting priority areas for building a competitive advantage, or by improving regional innovation eco-systems.

Our analysis ties into the comprehensive framework for analyzing the policy mix developed by Rogge & Reichardt (2016). Focusing on policy strategy and objectives, we address both consistency and coherence as key characteristics of effective policies. First, *consistency* is defined by Rogge & Reichardt as *"how well the elements of the policy mix are aligned with each other"*. While the European Commission expected consistency between RRP and S3 policies it was not trivial that RRPs tied into the existing smart specialization strategies of regions and countries given that the former were devised in a short timespan, under the pressure of the pandemic. Our analysis will show how the evaluation of the consistency of policies by an expert can be supported by algorithmic assessment. More specifically, we provide a proof of concept where the performance of the algorithm in judging the semantic similarity of pairs of policy objectives is benchmarked against the evaluation by an expert. The results show that the accuracy of the algorithm is sufficient to allow for reduced, i.e., more targeted, expert effort in assessing the consistency of large volumes of policy data. Second, our analysis indirectly addresses the notion of the *coherence* of underlying policy processes, defined by Rogge & Reichardt as *"synergistic and systematic policy making and implementation processes contributing towards the achievement of policy objectives"*. Coherence can be hindered by different policy frames or contradictory objectives. Conversely, coherence can be increased by developing overarching strategies and other efforts to create synergies between policy objectives (Kivimaa et al., 2023). The issue of coherence is pertinent to our empirical setting as RRP and S3 span different policy domains, that is, post-pandemic recovery and resilience (RRP) versus location-based innovation strategy (S3). Finally, as Rogge & Reichardt (2016) explain, the concept of coherence not only spans different policy domains but also applies across governance levels. Besides the RRP-S3 comparison (section 4.1), our analysis therefore also considers the federal-regional alignment of S3 strategies (section 4.2).

In sum, we assess the consistency of policy objectives through the presence of S3 priorities areas in the RRP as well as the coherence of policy processes - horizontally, i.e., across the RRP-S3 policy frames, and vertically, i.e., between S3 governance levels. We provide a methodological contribution by showing how this analysis can be supported by automated content analysis.

The results of our analysis show that the similarity measure produced by the semantic analysis algorithm is able to distinguish whether 2 policy measures are related (and to what extent), and that it serves as a valuable metric in the evaluation of policies' synergistic potential. Furthermore, we distil several takeaway messages on the combination of expert and algorithmic assessment, as guidelines for future efforts to successfully combine automated text analysis of policy documents. Finally, we show the versatility of the method by showing that it can be used to evaluate the alignment of policy measures at different levels of government, i.e., vertical policy coordination, and report on its robustness in other empirical settings.

The remainder of the paper is organized as follows. The next section provides a brief background on the Portuguese Recovery and Resilience Plan and its Smart Specialisation Strategies. Section 3 discusses the methodology, providing a brief technical explanation of the text analysis algorithm (Latent Semantic Analysis, LSA) used for establishing the semantic similarity between pairs of documents describing individual policy measures. This section also explains how the quality of the matches is assessed and outlines the overall approach of how the text analysis can be integrated into the workflow of a human expert. Section 4 discusses the results of the analyses of horizontal and vertical policy coordination. Section 5 provides a detailed reflection while section 6 concludes with the main takeaways and ideas for future research.

2. Background on the Portuguese Recovery and Resilience Plan and its Smart Specialisation Strategies

The European regional smart specialization policies illustrate how structural transformations of member states' economies are believed to be facilitated by tapping into synergies between various innovation-related activities. Foray *et al.* (2021) define a transformative activity as *"a collection of related innovation capacities and actions, all oriented towards a certain structural change"*. The post-pandemic recovery plans for Europe echo a similar ambition, aiming to absorb the impact of the COVID-19 pandemic while also making the member states more resilient against future shocks. The so-called *Recovery and Resilience Plans* (RRP) describe how the funds will be used, specifying milestones and targets to be achieved as well as reforms and investment projects to be implemented by 2026. The Portuguese RRP[2] is composed of 20 components (shown in the top panel of Table 1), further divided in 83 investments and 37 reforms, representing an investment amount of €16,644 million (República Portuguesa, 2021). The 20 components are thematic blocks of several measures and actions associated with investments and reforms, and grouped in three structural dimensions (resilience, green transition and digital transition). The Portuguese RRP, which received a positive assessment from the European Commission (2021b), has as a short-term objective the reactivation of demand due to the Covid-19 crisis and as a long-term goal the structural economic transformation of the country. The RRP funds come in addition to the contributions from the *EU cohesion policy*[3] and those from other EU sources, such as the

[2] We used the first version of the Portuguese RRP submitted on 22nd April 2023 to perform the analysis. A revised version of the plan has been submitted and was still under revision in 2023 (European Commission, 2023).
[3] The Cohesion policy receives funding from the *European Regional Development Fund* (ERDF), the *European Social Fund* (ESF) and the *Cohesion Fund* (CF).

European Agricultural Fund for Rural Development (EAFRD), the *European Maritime and Fisheries Fund* (EMFF), the *InvestEU* and *Horizon Europe* missions and partnerships. Therefore, EU Member States may benefit from synergies and complementarities between these different financing schemes.

A particular source of potential synergies with the RRP comes from the *Smart Specialisation Strategy* (S3) approach. S3 refers to a governance model using a place-based approach, involving stakeholders in the policy-making process (Foray *et al.*, 2009). It aims to make regional and national innovation policies more effective by helping the identification of innovation priorities for investment. While having an S3 in place was a so-called *ex-ante conditionality* in the programming period 2014-2020 and an *enabling condition* in 2021-2027 to be allowed to spend the *European Structural and Investment Funds* (ESIF) on Research and Innovation, the European Commission has encouraged Member States to use other policies and sources of funding to implement the strategies (European Commission, 2012). A case in point is the Next Generation EU, in particular the RRPs, as a noteworthy example of a policy that can provide additional financing of S3. Portugal has adopted a multi-level approach to Smart Specialisation that includes both a national and seven regional strategies, with the latter corresponding to five continental regions and two autonomous archipelagos (Azores, Madeira). The second panel in Table 1 shows the innovation priorities for the programming period 2021-2027.[4] The potential for alignment between the RRP and the Portuguese S3 policies is arguably present. The Portuguese RRP explicitly states the ambition to align R&I projects in the RRP with S3 priorities (República Portuguesa, 2021). Furthermore, the RRP mentions that the *Development and Cohesion Agency* is part of the committee in charge of the technical coordination and monitoring of the RRP. This agency coordinates the Regional Development Policy, ensures the general coordination of the European Structural and Investment Funds (ESIF), and, importantly, also participated in the working groups to discuss the S3 national priorities for 2021-2027.

The European Commission already conducted a study of the Portuguese RRP and its annexes based on a detailed “manual” text analysis, i.e. conducted by a human expert, to detect investments that can enhance Research, Development and Innovation activities and to improve the regional innovation eco-systems (Marques Santos, 2021). This mapping was then matched by an expert with the national and regional S3 innovation priorities for 2021-2027. In this paper we do a similar exercise, again assessing the RRP-S3 alignment for Portugal, but instead of relying on expert analysis alone the approach integrates semantic text analysis. More specifically, we use the latent semantic analysis (LSA) algorithm to compare the documents that describe, respectively, the Portuguese RRP components and those containing its national and S3 priorities in 2021-2027. The method is explained in more detail in the next section.

3. Method

3.1 Background on Latent Semantic Analysis

[4] The smart specialization priorities in Portugal are documented in CCDR-Alentejo (2021), CCDR-LVT (2020), CCDR-Centro (2021), CCDR-Norte (2021), ANI and Quaternaire Portugal (2021), as well as information provided to the authors by regional managing authorities in Algarve, Azores and Madeira.

The goal of the analysis is to compare 2 sets of policy documents (one corresponding to RRP, the other to S3) to obtain a measure of the semantic similarity of pairs of text documents. However, determining whether documents have a similar meaning is challenging. The first problem is that the same word can have different meanings in a different context. For example, a term like *"modalities"* carries a different meaning when used in the context of a policy that aims for intelligent urban traffic flows (where it may refer to combining different modes of transport like car, bus and bike) versus when it is used in the context of lifelong learning (where it may refer to online versus face-to-face modes of learning). The second, opposite problem is that different words can have the same meaning across contexts. For example, the words *"shift"* and *"transition"* can be considered to essentially mean the same, whether used in the context of greening road transport or in the context of firms switching to hydrogen in the industrial value chain.

These problems imply that it would be unreliable to determine the similarity between documents by merely relying on exact matches with pre-defined keywords as a measure of similarity. In addition, a manual approach of reading and hand coding documents becomes prohibitive for dozens, let alone hundreds, of documents. For these reasons algorithmic approaches are seen as a promising alternative, in particular those text analysis algorithms capable of determining *semantic* similarity between documents, such as Latent Semantic Analysis (LSA). Besides offering the benefit of scalability, a key strength of LSA is that it makes use of word co-occurrences but does not rely on matching specific keywords to documents (Valdez et al., 2018; Schwarz, 2019). Also, it does not infer the meaning of words based on one document alone but looks across documents in the text corpus. LSA has become increasingly popular and has been used in a variety of contexts, such as the identification of R&D partners for firms by analysing firms' patent documents (Park et al., 2015), the analysis of the coherence of patient narratives in a medical context (Vrana et al., 2019), or even early attempts at automated scoring of student essays (Landauer et al., 2003). The result of the semantic text analysis is a single number for each RRP measure that indicates how semantically similar it is to each S3 priority, allowing to assess the horizontal coordination between both policies (section 4.1). Cosine similarities are also calculated to assess the "vertical" alignment of national and regional S3 priorities (section 4.2). The details on the calculation of the similarity between two policy documents and the required pre-processing of the input documents are included in Appendix 1.

3.2 Evaluation of text analysis results

A key message of this study is that automated text analysis needs to be complemented with expert input in order to contextualize results and maximize its usefulness. Figure 1 shows the inputs and outputs of the text analysis and how these integrate with expert assessment. The initial step consists of the selection of the source texts that will be assessed by the text analysis algorithm and the expert. The goal is to construct a "relevant sample", i.e., a text base that provides a representative description of the phenomenon under study, in our case the alignment of two policies. After this initial *sampling validity* step, the text analysis is carried out[5], comparing the semantic similarity between the components of the

[5] We specified 10 components for the Singular Value Decomposition (see section 3.1) performed by the LSA algorithm.

RRP and the S3 priorities.

[Figure 1]

As explained in section 3.1, the semantic text analysis yields a measure of the similarity between two documents. Building on this measure, we adopt the following procedure in order to assess the quality of the matches.

Strength of the match. We impose that the cosine similarity score must exceed a minimum threshold.[6] More specifically, we require that the cosine similarity of a given RRP-S3 comparison is higher than the average RRP-S3 similarity plus 1 standard deviation to identify a strong match.

Algorithmic robustness of the match. We verify the consistency of the results across different parameterizations of the algorithm. More specifically, we run separate analyses using 10 and 20 components for the Singular Value Decomposition (see section 3.1), the number of dimensions used to capture the variation between documents. We assess whether there is any overlap in the top 2 matches when using 10 or 20 components, respectively and consider the match sufficiently robust if one or both of the top-2 matches is the same.[7]

Assessment of accuracy, precision and recall. We assess the classification performance of the algorithm in terms of its *accuracy*, *precision* and *recall*. These indicators are commonly used in research that relies on text analysis to classify documents (e.g. Belderbos et al., 2017). In particular, we rely on an ex-ante assessment by a human rater — an expert in EU policy and very familiar with the RRP and S3 policies — who evaluated each RRP measure as having a *low*, *medium* or *high* matching potential with one or more of the smart specialization priorities. The following criteria were used by the expert to evaluate whether a component of the RRP plan matches with an S3 priority (see Figure 1):

- Whether it contains an explicit focus on ***R&D*** (bringing it into the S3 policy realm);
- Whether it aligns ***thematically*** with an S3 priority;
- Whether it links to an S3 priority either ***directly or indirectly***[8];
- Whether it specifies an ***investment and beneficiaries*** that target the S3 priority.

These criteria reflect the potential for complementarities between elements of the RRP plan and the national/regional S3 priorities. If the expert assessed an RRP measure as meeting all 4 criteria, it was flagged as having a *high* expectation to match with an S3 priority. If none of the criteria were met, it was marked with *low*, and the remaining cases as *medium*. These labels allow evaluating the classification performance of the LSA algorithm by distinguishing between 4 categories for each RRP-S3 match:

1) *True Positive* (TP): high RRP-S3 cosine similarity and high expert expectation of a match;[9,10]

---

[6] Note that there isn't a specific cut-off value for the cosine similarity that indicates unambiguously whether there is a link between the RRP and the S3 priority. The similarity between two documents is a matter of degree and a value of, say, 0.75 is essentially arbitrary. A threshold that accounts for the specificity of the given text base should rely on the entire distribution of cosine similarities.

[7] An alternative criterion could be that the *best match is identical* when using either 10 or 20 components for the Singular Value Decomposition. However, this strict requirement ignores that an RRP component may match with more than one S3 priority, which is a valid outcome. In other words, there needn't be only a single best match and considering the top-2 matches allows for some flexibility in this regard.

[8] An RRP-S3 link is considered to be indirect if the RRP component relates to a so-called *transversal* S3 priority, which cuts across multiple S3 priority domains.

2) *False Positive* (FP): high RRP-S3 cosine similarity but low expert expectation of a match;
3) *True Negative* (TN): low RRP-S3 cosine similarity and low expert expectation of a match;
4) *False Negative* (FN): low RRP-S3 cosine similarity but high expert expectation of a match.

The 4 categories *True Positive*, *True Negative*, *False Positive* and *False Negative* allow constructing 3 diagnostic indicators that are informative with respect to the algorithm's ability to produce accurate matches. First, considering all RRP measures for which there was a clear expectation about whether or not they would match with an S3 priority (i.e. those with a *low* or *high* expectation of a match), ***accuracy*** is the share for which the algorithm was correct. In other words, for all RRP measures that could be classified as having either a strong or weak link with an S3 priority, how often did the algorithm yield a high, respectively low, cosine similarity with an S3 priority? Formally, *accuracy* is expressed as *(TP+TN)/(TP+FP+FN+TN)*. Second, ***precision*** tells us, considering the RRP measures that were matched to an S3 priority, how many the algorithm classified correctly. In other words, out of the RRP measures that the algorithm considers to be very similar to an S3 priority (=TP+FP), for how many did we actually expect a match (TP)? Formally, *precision* equals *TP/(TP+FP)*. Finally, the ***recall*** indicator tells us, considering the RRP measures that should have been matched to an S3 priority, how many the algorithm classified correctly. In other words, out of the RRP measures that the human rater considers to be very similar to an S3 priority (=TP+FN), for how many did the algorithm produce a match (TP)? Formally, *recall* equals *TP/(TP+FN)*.

Note that it is critical that the criteria used by the expert in the ex-ante assessment of a match are made explicit as this permits resolving any discrepancies between the text analysis and the expert assessments, shown as the concluding step of the procedure in Figure 1. For example, the results may show a strong match between an RRP component and an S3 priority in the text analysis although the expert flagged this RRP component as having a *low* ex-ante linkage with smart specialization. This discrepancy can be traced back to the individual criteria in order to reveal that, for example, the text analysis picked up on the thematic RRP-S3 alignment but may have missed the lack of an explicit R&D focus in the RRP component. The expert may then decide to adjust her own assessment and follow the match found by the text analysis, which makes sense if the RRP *does* refer to technological innovation that the expert may have overlooked. Conversely, the expert may decide that the text analysis produced a 'false positive' and stick with her judgment. In section 4.1 we provide various examples of how differences in algorithmic and expert assessment can be reconciled.

4 Results

4.1 Horizontal Policy Coordination

We start by analysing the alignment of the Portuguese RRP and S3 based on the previously outlined evaluation logic. Table 1 mentions the number of documents describing the Portuguese RRP and its S3

[9] "high similarity" means a cosine similarity above the threshold value i.e. the average similarity plus 1 s.d.
[10] In the assessment of the *True Positives*, it is required that the match is algorithmically robust. This requirement implies that if an RRP-S3 match shows a high similarity but is not algorithmically robust, it becomes a FN rather than a TP. The empirical results indicate this is rare.

strategies, as well as their average word length. There are 20 components in the Portuguese Recovery and Resilience Plan with the average length of the description equal to about 400 words. The national and regional smart specialisation strategies contain 4 (Azores) to 8 (Lisbon, Norte, Alentejo) priorities with the average length of the description equal to 355 words.[11]

[Table 1]

The tables in Appendix 2 illustrate the semantic relations between words in the input documents. Table 5 in Appendix 2 shows the national smart specialisation priorities as well as the ones for the Lisbon region, joint with an extract from each description and its total word count. Despite the conciseness of the documents, they capture the gist of the priorities well while one can also note the different emphasis put in similar priorities of different authorities. For example, the national S3 priority domain *Health, Biotechnology and Food* explicitly mentions these 3 focal points, while the *Health* priority in the Lisbon region also mentions "ageing" as a point of attention. Recall that the LSA algorithm does not look for exact matches of certain words to determine the similarity of two documents (section 3.1). Rather, it is capable to identify semantic relationships between words by considering co-occurrences of words in the entire document. Table 6 in Appendix 2 illustrates this semantic similarity for the RRP components: it shows the most frequently occurring word in the document describing the component along with words that have been identified by the LSA algorithm as being highly similar.[12] So, for example, the RRP component *The National Health Service* unsurprisingly has "health" as the most frequent word in its description. Based on the comparisons of documents, the LSA algorithm is able to determine that words like "well-being", "care" and "ageing" are often used in a health context within the text corpus. Even words that are syntactically very different and that aren't even remotely synonyms, such as "telemonitoring", are found to be indicative of a health context.

We discuss the results of the comparison between the RRP and the *national* S3 strategy following the procedure in section 3. The detailed results are shown in Table 7 in Appendix 2: it contains the 20 RRP components of Portugal and the best matching national S3 priority i.e. the one that yields the highest cosine similarity.[13]

For example, for the 1st RRP component *The National Health Service*, the national S3 priority domain with the highest match is *Health, Biotechnology and Food*, with the cosine similarity of the 2 documents equal to 0.70. The 2nd best match (not shown in the table) is with *Major Natural and Emerging Assets - Space and Earth Observation Technologies*, but with a much lower cosine similarity of 0.04. The latter value is below the threshold, which is equal to 0.35 i.e. 1 standard deviation (0.20) above the average cosine similarity (0.15) of all pairwise comparisons between the RRP components and all national S3 priorities.

[11] The length of the descriptions varies, as it depends on the respective source documents that differ in their verbosity. Note that the LSA-algorithm reduces the weight of terms that occur frequently (section 3.1), which (partially) addresses the issue of higher word frequencies due to the greater length of some documents.
[12] These highly similar words are subsets from the top-10 words that yield the highest cosine similarities with the most frequent word in the RRP component, based on the word-component matrix generated by the LSA algorithm.
[13] These results are based on an analysis that uses 10 components for LSA Singular Value Decomposition. Detailed results, including those using a higher number of components, can be consulted in the files referenced in Appendix 2. The reported results are robust to variations in algorithmic parameterizations.

Analogously, the 2nd RRP component *Housing* yields *Large Natural and Emerging Assets – Sea* as the best matching national S3 priority. This result doesn't seem intuitive, which is confirmed by the cosine similarity value (0.34), which is below the threshold. Since the best match doesn't jump over the hurdle for a good match, this RRP component can be considered as not aligned with the national S3 strategy, marked by "—" in the table.

In total, 15 of the 20 RRP components (75%) yield a match with a national S3 priority, i.e., with a cosine similarity above the threshold. Note that this value mustn't be interpreted as "higher is better", since an outcome where *all* RRP components are matched to an S3 priority may rather indicate a lack of discriminatory power of the algorithm. The share of matched RRP components does indicate that the Recovery and Resilience Plan is fairly well aligned with the national S3 strategy.

The 4th column of the table shows the algorithmic robustness of the results. As explained in section 3, we consider an RRP-S3 match (when using 10 components for the singular value decomposition in the LSA algorithm) as robust if either of the top-2 matches coincides with one of the top-2 matches when using 20 components (C=20). This is generally the case: only RRP component *Infrastructure* fails this test. Its top 2 matches are *Territory, Creativity and Brands* (cosine=0.37, just above the threshold) and *Large Natural and Emerging Assets – Sea* (cosine=0.32, below the threshold, not shown in the table). When using C=20, the top 2 matches for *Infrastructure* are *Major Natural and Emerging Assets - Space and Earth Observation Technologies* (cosine=0.15) and *Economy 4.0 AND Digital KET* (cosine=0.13), respectively, which don't overlap with the top 2 matches for C=10.[14] The algorithmic robustness is only evaluated for matches above the threshold, so not for weak matches. For reliability, a higher value is of course always preferred. For the 15 RRP components that yield a sufficiently strong match with a national S3 priority, 14 of them are robust.

The final column of Table 7 in Appendix 2 shows the classification performance of the LSA algorithm, which confronts the matches of the algorithm with those of a human rater, the latter shown in the column "*Expectation of match*". As explained in section 3, this column contains the rater's expectation how likely each RRP component is to match with an S3 priority domain i.e. *low*, *medium* or *high*. Recall that the comparison with the human rater's judgment allows classifying the strong and robust matches of the algorithm into different categories: *True Positive* (TP: a match that was also expected by the human rater), *True Negative* (TN: no match, as expected by the human rater), *False Positive* (FP: a match, against expectation) or *False Negative* (FN: no match, against expectation).[15] As an example, RRP component 11, *Decarbonisation of Industry*, is matched with S3 priority *Circular Economy, Energy Transition and Decarbonisation*, as expected, and is therefore classified as a *True Positive*. Conversely, RRP component 3, *Social Responses*, was not expected to be matched to any S3 priority but yields a strong match with *Health , Biotechnology and Food*, and is therefore classified as a *False Positive*. Overall, the classification performance of the LSA algorithm for matching the RRP components with the national S3 strategy can be considered fairly high, with an *accuracy* of 69% (considering both TP and TN) and a *recall* of 100% i.e. it matches all RRP components it should have matched to an S3 priority according to the human rater. Conversely, the algorithm achieves a *precision* of only 50% i.e. the

[14] The results for C=20 are not shown in the table due to space limitations but are available in the full results file, see Appendix 2.

[15] The assessment of the classification performance of the algorithm is restricted to those cases for which the human rater has a clear expectation about the RRP-S3 alignment (*high* or *low*).

algorithm was right in only half of the cases when it declared a match. Several comments help understand this result. First, note that there is an inherent trade-off between *precision* and *recall*, since greater success in "catching" all possible RRP-S3 matches (and thus a higher *recall*) increases the likelihood of false positives, reducing *precision*. Second, the classification performance depends on the available information in the text documents that are fed to the algorithm, and these may contain more limited information than the sources the expert has access to or implicitly takes into account. Finally, the assessments of a human rater are not necessarily without error. A more extensive approach could therefore consist of having two raters evaluating the matching potential of the RRP components independently and retain only the cases where they reached the same conclusion in order to minimize human error. We return to these points in the conclusions.

We also report the results of the comparison of the Portuguese RRP with one of the regional smart specialisation strategies, namely the one of the Lisbon region in Table 8 in Appendix 2. The interpretation of the results follows the same logic as for the analysis of the RRP and the national S3 strategy, so here we focus on highlighting a few examples that clarify how the ex-post assessment of the text analysis results by the expert (as the final step in the procedure, see Figure 1) provides valuable insights.

A first example illustrates how the text analysis yields a strong and robust match for RRP component 1 on *The National Health Service* and S3 priority *Health* (cosine similarity=0.78, well above the threshold of 0.47). However, ex-ante the expert did not expect a match, which means the match is classified as a *False Positive*. Inspection of the results by the expert reveals that, despite the thematic alignment on the improvement of primary health care, the low expectation of a match was driven by a lack of focus on R&D and innovation (expert criterion 1, see section 3). The RRP component indeed mentions innovation only once: *"This component aims to strengthen the capacity of the National Health Service (NHS) to respond to the country's demographic and epidemiological changes, therapeutic and technological* ***innovation****, rising health costs and the expectations of a more informed and demanding society […]"*. The expert therefore decided to stand her ground and consider the RRP initiative as insufficiently targeted towards R&D and innovation in order to consider it closely aligned with the Lisbon S3 Health-priority.

A second example illustrating a discrepancy between the ex-ante human expert assessment and the algorithm's result concerns RRP component 11 on the *Decarbonization of Industry*. The expert's high expectation of a match with the Lisbon S3 strategy is based on the existence of an indirect link (expert criterion 3, see section 3), as this RRP component relates to a *transversal* S3 priority (Marques Santos, 2021). However, because the transversal smart specialization priorities were not included as input for the algorithm, it can only identify matches with non-transversal, domain-specific S3 priorities and none of these pass the cosine similarity threshold, resulting in a *False Negative* classification.

A final example concerns RRP component 4, *Culture*. Ex-ante the expert did not have strong expectations of a match with a regional S3 priority, while the text analysis produces a strong and robust match with Lisbon's S3 priority on *Creative and Cultural Industries*, with a high cosine similarity value of 0.80. Feedback from the expert reveals that she took into account the beneficiaries of the investments associated with the RRP component (expert criterion 4, see section 3), whom are located both within and outside of the Lisbon region, and this misalignment resulted in her assessment of the link between the RRP component and this *regional* S3 priority as *medium*. Similar to the previous example, the cause for the discrepancy is that certain information (in this case on the investment targets) is not present in the textual descriptions of the RRP components that were fed to the algorithm.

Besides the two preceding analyses, the Recovery and Resilience Plan of Portugal was also compared with the remaining 6 regional smart specialization strategies.[16] The results of the 8 matching exercises are summarized in Table 2. On average, somewhat more than half (57%) of the RRP is aligned with the different S3 strategies, although this unweighted average takes into account economically smaller regions like the Azores (30%) and the Algarve (45%) to the same extent as the Lisbon region, Norte and Centro, where alignment is higher, at 60%. The algorithmic reliability is generally very high (98%), which indicates that the quality of the match does not depend on the precise parametrization of the algorithm. We see more variance in the classification performance of the algorithm (see the panel *"based on 4 expert criteria"*) across the regions, with generally high scores for *recall* and — logically, as there is a trade-off between the 2 metrics — lower scores for *precision*. Note that the classification performance depends on the amount of information used by the expert: the more contextual information she takes into account that is unavailable to the algorithm, the lower the *precision*. To further illustrate this point, consider the righthand panel in Table 2. Here, the RRP-S3 matching was done based on thematic alignment only rather than all 4 criteria (Figure 1), which puts the expert and the algorithm on more equal footing. In this case, we obtain a higher share of *True Positives* and hence the precision of the algorithm goes up. Note it would be misguided to try and maximize classification performance without considering the underlying reasons. Rather, any classification performance aids the expert in gaining more in-depth understanding of *why* her assessment coincides with the algorithmic results or not.

[Table 2]

As a final robustness check for the analysis on horizontal policy coordination, we applied the same approach to the RRP and regional S3 strategies of Belgium. The key difference with the Portuguese case is that both the RRP and S3 policies in Belgium are fully regionalized, which implies that the analysis of RRP-S3 alignment can be conducted separately by region. This analysis by and large confirms the insights from the Portuguese case, indicating that the approach is generalizable to other countries. For brevity, the results for Belgium are not discussed in the paper but are available in the online appendix.[17]

4.2 Vertical Policy Coordination

To show the versatility of the method, we also use it to analyse vertical policy coordination. In a federal regime where regional governments have substantial autonomy, the alignment between different levels of government creates yet another coordination challenge besides the one between different policy realms (Griessen & Braun, 2008; Braun, 2008; Bolleyer & Borzel, 2010; Peters, 2018). For this purpose, we zoom in on the smart specialisation strategy of Portugal and, using the same algorithmic approach, analyse the alignment of the national smart specialisation strategy with the ones of the regions. This exercise complements the previous analysis by showing how the approach can be used to analyse coordination of policies at different governmental levels within a single policy frame, in this case smart specialisation.

[16] Detailed results for the other regions are referenced in the files in Appendix 3.
[17] We can provide these results to referees if needed.

We begin by taking the perspective of the national smart specialisation strategy and analyse for each national priority to what extent it is reflected in the 7 regional smart specialisation strategies. Table 3 shows that, for example, the national S3 priority *Territory, Creativity and Brands* matches with a similar priority in each of the 7 regions while the national priority *Large Natural and Emerging Assets – Sea* matches with 5 regions. The results at the level of individual S3 priorities[18] show that this S3 priority has no counterpart in, rather intuitively, the Lisbon and Centro regions. Matches are determined in an analogous fashion as in the analysis of the RRP-S3 alignment: the cosine similarity of a national and regional S3 priority is calculated and if it is greater than the benchmark, the regional priority is considered a match with the national priority. The benchmark is determined as in the RRP-S3 analysis i.e. 1 standard deviation above the average cosine similarity for all pairwise (national, regional) S3 priority comparisons. Conditional on crossing that threshold, the average cosine similarity for national and regional S3 priorities is 0.73 (bottom row, Table 3). Note that the cosine similarity complements the information on the national-regional alignment of S3 strategies relative to the mere count of matching regions. For example, while the national S3 priority *Major Natural and Emerging Assets - Space and Earth Observation Technologies* is present in only 4 of the 7 regions, we see that *conditional on a match* the alignment is quite strong (cosine=0.74). The detailed matching results (not reported here) show that there is a strong match especially with the regional strategies in Norte (0.81), Centro (0.86) and Azores (0.70, linked to its radio astronomy and space infrastructure).

[Table 3]

Finally, we consider the opposite direction i.e. from the perspective of a given region we verify to what extent its regional smart specialisation priorities are reflected in the national S3 strategy. We evaluate the results, shown in Table 4, using 2 indicators. First, the share of S3 priorities of the region that yields a match with a national S3 priority shows that, on average, 88% of regional priorities has a national counterpart, ranging from 67% for Algarve (where 2 regional priorities don't yield a match) to complete alignment for the Alentejo, Centro, Madeira and Norte regions. Second, analogous to the national-regional analysis, we consider the cosine similarity of matching national priorities, which gives an indication of the strength of alignment, conditional on a match. The average cosine similarity is 0.71, which - as noted before - cannot be interpreted in an absolute sense but given that a cosine value ranges from -1 to +1, the similarity between matching regional and national priorities can be considered rather high. Also here, note that the measures (the share of S3 priorities with a match and the average cosine similarity conditional on a match) complement each other. For example, while all regional S3 priorities in Alentejo match with a national priority, the average strength of these matches is fairly modest (cosine 0.61, with the threshold for a match equal to 0.49). Conversely, Madeira also has full alignment with the national strategy but with a relatively much stronger match, on average (cosine 0.87).

[Table 4]

[18] These detailed results are not are reported here, but are available in the files referenced in Appendix 2.

5 Discussion

The objectives of this paper were to explore the use of automated text analysis to assess policy coordination and to learn how it can complement expert assessment. Using a case study, we provide a proof of concept of how automated text analysis can be used to analyse the alignment between different policies and how the technique can be integrated with expert assessment. We believe that the approach can be useful for both policy analysts and policy makers. Policy analysts in academia and government are the primary audience, given the methodological focus of the paper and its applicability to other policy areas. The results of an analysis such as the one discussed in section 4 are rather targeted at policy makers at regional, national and European level as it will help them to assess the consistency of policy objectives and the coherence of the underlying policy processes. Based on the empirical analysis, we now draw lessons about the necessary conditions for the successful implementation of automated text analysis as well as its potential in other empirical settings.

The first conclusion relates to the algorithmic assessment of the semantic similarity between policy measures. The LSA algorithm delivers on the promise to analyse large volumes of data much faster than a human analyst ever could. This scalability is a marked advantage and opens the door to applications where a large number of policy instruments need to be assessed, such as in the European industrial policy (Terzi et al., 2023). Furthermore, given that there have been few suggestions in the literature for indicators that can measure policy coordination (Tamtik, 2017), a key strength of the approach is that the cosine similarity measure produced by the algorithm performs well in terms of determining to what extent 2 policy measures are aligned. However, as the analysis in section 4.1 showed, a high score on the similarity measure should not be taken at face value. Even though the LSA algorithm reduces the likelihood of false positives relative to a more straightforward approach that relies on matching specific keywords (see section 3.1), it may not pick up on certain policy dimensions and a high semantic similarity does not equate alignment. Hence, while the cosine similarity measure is not a definitive assessment of the (lack of) consistency of policy measures, it is a valuable starting point for the evaluation of policies' synergistic potential: for complementarities between policy initiatives to be realized, they arguably need to exhibit some degree of relatedness to allow for fruitful interaction. The method presented in this paper therefore fits in a policy mix perspective (Flanagan et al., 2011; Diez 2002; Aranguren et al. 2014) by identifying "thin crossing points" between policies. In future work, we envisage that the measure of semantic similarity can be put to use in various applications of research evaluation. In particular, it can serve as the input for an analysis that considers the procedural mechanisms underlying the analysed policies to understand why they are (not) coordinated (Cejudo & Michel, 2022), contributing to an assessment of the coherence of underlying policy processes. A specific example in which the approach could be leveraged is to study the extent to which subsidized projects align with the original policy objectives. Another example using automated text analysis would be to assess to what extent distinct policies put in motion the same macro-processes to achieve transformative change, e.g. using Kivimaa et al.'s (2023) taxonomy.[19] While these are exciting ways to apply automated text analysis to policy documents, much work remains to be done with respect to developing additional indicators for higher-order concepts like consistency and coherence.

Second, we identify a number of potential pitfalls that must be taken account, responding to Grimmer & Stewart's (2013) call for caution when using automated text analysis. We articulate hands-on advice for the combination of expert and algorithmic assessment that may serve as a guideline for future efforts, not limited to the domain of research and innovation. In our discussion, we point out the limitations of our work that can be addressed in future efforts. First, semantic analysis using algorithms like LSA offers the most potential if thoughtfully embedded in the workflow of experts (see Figure 1). A fully automated approach without expert input is unlikely to yield results that are sufficiently reliable to serve as an input for policy evaluation. A first reason is that algorithms are not infallible: deducing the semantic meaning of a text remains a tough problem. An inherent restriction on the algorithm's performance are the limitations of the textual descriptions used as input due to their varying length or richness. As a result, the assessment of the consistency of two policies may typically be limited to thematic alignment while a more in-depth assessment would also take into account whether 2 policies operate within the same geographic area, play out over the same time horizon, target the same beneficiaries, are financed from common resources, etc. A second, related issue is that not all information available to experts may be readily convertible to text documents that can be used for the analysis due to confidentiality issues or difficulties in converting quantitative or financial information into full text. These considerations suggest that the most promising approach for using automated text analysis is to *support* the work of the expert rather than aiming to replace the expert entirely.

A second attention point relates to what we referred to as *sampling validity*. The automated text analysis can evidently only process the text base that is it is given. Consequently, differences in the text base that serves as the input for the text analysis on the one hand and the information used by the expert on the other hand will increase the discrepancy between the two assessments. More specifically, if the text base used for the text analysis consists of merely short descriptions of the RRP components while the expert has access to additional data, for example on investments and their beneficiaries (some of that information may even be tacit), the automated text analysis cannot take this into account and discrepancies with the expert's assessment are to be expected. Hence, it is important to include all available information in the text base that is used as input for the text analysis and to verify that it allows evaluating all criteria taken into consideration by the expert so the two approaches are on a level playing field. That said, even if all available information is included in the input texts, it may not be trivial for a text analysis algorithm to pick up on logical conditions that can be readily evaluated by a human expert. For example, if a policy targets beneficiaries active in a certain domain and region, then a semantic algorithm may consider it consistent with another policy that targets the same domain but different beneficiaries A solution in such cases is to put additional structure on the data by preprocessing the data so the algorithm can be fed with policy measures that are known to have matching eligibility conditions.[20]

---

[19] Note that such alignment would not be trivial in the case of RRP and smart specialization strategies since the latter focus by design on regional strengths and may therefore be less geared towards regime destabilization whereas the RRP may focus more on promoting niches.

[20] For example, in the case of a country like Belgium, RRP and S3 policies are fully regionalized by design, so they can be analyzed separately by region.

Third, it is key to assess the robustness of the automated text analysis as well as the expert assessment. For the automated text analysis, it is important to understand that it can never be "fully automatic" since each algorithm requires parameters to be set by the researcher. Examples of such choices include the cosine similarity threshold used for defining a "strong" match or parameters such as the number of components for reducing the dimensionality of the data. It is critical to report such choices transparently and to carry out robustness checks to establish trust in the results, as we have done in our analysis. Conversely, to increase the reliability of the ex-ante expert assessment, it is recommended to have multiple raters independently evaluate the correspondence between policy measures. By verifying inter-rater agreement and reconciling any differences between assessments by the raters, the role of human error can be reduced as a source of discrepancies between semantic text analysis and expert judgment. Nevertheless, besides the added cost when using multiple experts, combining multiple experts' judgments may still result in biased results when using weights that don't reflect their respective quality (Hammitt & Zhang, 2013). The commonly used practice of assigning an equal weight to each expert assumes each one has the same amount of relevant expertise and devoted an equal effort to the task, an assumption that is hard to verify and that may easily be violated in real settings. Conversely, algorithmic approaches do not suffer from such bias, provided the choices for the parameters that govern their operation are transparently reported, as mentioned above. An algorithmic approach to policy analysis can therefore deliver on objectivity and reproducibility in a way that is harder to realize with expert assessment alone.

Finally, given the swift developments in natural language processing, we reflect on our choice for Latent Semantic Analysis as the algorithm underlying the approach. The method was chosen due to the empirical setting, in which we compare sets of policy documents, which are relatively short pieces of text that are less amenable to an analysis using word embedding algorithms, like *word2vec*, that are based on deep learning, which has given rise to large language models (LLMs). Such LLMs that have been pre-trained using everyday language may not be well suited to the specific language use in policy documents, which calls for retraining the models. However, the relatively small number and the limited length of the descriptions of policy objectives may make them unattractive to train a LLM, and algorithms such as LSA have been shown to perform better in such cases (e.g. Altszyler et al., 2016). That said, LLMs may offer potential for a more comprehensive assessment by taking into account geography, time horizon or beneficiaries, but whether they can provide a better measure of the consistency of policies is an unresolved issue at this point (Shen et al., 2023). A step forward may consist of combining the strengths of different methods, for example by using word embeddings to adjust the weighting schemes for the input matrix of the Latent Semantic Analysis method (e.g. Al-Sabahi et al., 2018).

6 Conclusion

Provided that the above considerations are taken into account, we believe that incorporating automated semantic analysis into the expert's workflow is valuable. The advantages can be summarized as follows. First, while a fully algorithmic assessment of the alignment of policies seems too ambitious, the approach offers increased efficiency. Compared to a manual approach, the expert may initially with be satisfied with a quick assessment of the alignment potential of two policy measures and focus in-depth analysis on those cases where the textual analysis seems to disagree with expert's own assessment, especially those in the "false positives" category. Such a procedure has the potential to reduce the workload of the expert, especially when large volumes of documents need to be processed. A second advantage is that it provides the expert with a sounding board to verify one's judgment, thereby increasing the reliability of the exercise. Third, the semantic text analysis complements the expert's assessment through the cosine similarity, which provides a quantitative measure of the strength of the match that is not provided by purely descriptive techniques such as word-clouds.

In summary, the ex-post validation of the results by juxtaposing the text analysis matches and the ex-ante expert assessments offers a more robust end result than either approach is likely to provide on its own. We believe that the combination of large textual datasets describing policies and the implementation of algorithms for semantic text analysis algorithms in many software packages is a promising avenue for policy analysis. Given its potential, the integration of such methods into the policy analyst's toolbox is likely to attract much interest in the near future, and we hope that the above guidelines and considerations will support these efforts.

References

Al-Sabahi, K., Zuping, Z., & Kang, Y. (2018). Latent semantic analysis approach for document summarization based on word embeddings. *arXiv* preprint arXiv:1807.02748.

Altszyler, E., Sigman, M., Ribeiro, S., & Slezak, D. F. (2016). Comparative study of LSA vs Word2vec embeddings in small corpora: a case study in dreams database. *arXiv* preprint arXiv:1610.01520.

ANI and Quaternaire Portugal (2021). Transferência de Tecnologia em Domínios Estratégicos Nacionais: Súmula dos Resultados, Workshop Final, Webinar 27 de May de 2021.

Belderbos, R., Grabowska, M., Leten, B., Kelchtermans, S., & Ugur, N. (2017). On the use of computer-aided text analysis in international business research. *Global Strategy Journal*, 7(3), 312-331.

Braun, D. (2008). Lessons on the political coordination of knowledge and innovation policies. *Science and Public Policy*, 35(4), 289–298. https://doi.org/10.3152/030234208X310347

Bolleyer, N., & Börzel, T. A. (2010). Non-hierarchical policy coordination in multilevel systems. *European Political Science Review*, 2(2), 157–185. https://doi.org/10.1017/S175577391000007X

Boston, J. (1992). The Problems of Policy Coordination: The New Zealand Experience. *Governance*, 5(1), 88–103. https://doi.org/10.1111/j.1468-0491.1992.tb00030.x

CCDR-Alentejo (2021). Estratégia Regional de Especialização Inteligente, EREI2030: uma (nova) década para coconstruir região e desafiar o future, Alentejo 2030, Comissão de Coordenação de Desenvolvimento Regional do Alentejo. https://www.ccdr-a.gov.pt/estrategia-regional-de-especializacao-inteligente-do-alentejo-2030/

CCDR-Centro (2021). Estratégia Regional de Especialização Inteligente do Centro: Revisão para o período 2021-2027, RIS3 Centro, Comissão de Coordenação de Desenvolvimento Regional do Centro, April 2021. https://ris3.ccdrc.pt/

CCDR-LVT (2020). Estratégia Regional de Especialização Inteligente de Lisboa 2021-2027: RIS3 Lisboa, Comissão de Coordenação de Desenvolvimento Regional de Lisboa e Vale do Tejo, July 2020. https://www.ccdr-lvt.pt/wp-content/uploads/2023/03/RIS3_Lisboa_2030_v1_4.pdf.pdf

CCDR-Norte (2021). Estratégia de Desenvolvimento do Norte para Periodo de Programacão 2021-27 das Politicas da Uniao Europeia, Comissão de Coordenação e Desenvolvimento Regional do Norte. https://www.ccdr-n.pt/storage/app/media/2021/CCDRN%202030-compactado.pdf

Cejudo, G. M., & Michel, C. L. (2021). *Instruments for Policy Integration: How Policy Mixes Work Together. SAGE Open*, 11(3), 21582440211032160. https://doi.org/10.1177/21582440211032161

Chakraborty, B., & Bhattacharjee, T. (2020). A review on textual analysis of corporate disclosure according to the evolution of different automated methods. *Journal of Financial Reporting and Accounting*, Vol. 18 No. 4, pp. 757-777.

European Commission (2012). Guide to Research and Innovation Strategies for Smart Specialisation, Luxembourg: Publications Office of the European Union.

European Commission (2020). Directorate-General for Research and Innovation. Dixson-Declève, S., Balland, P., Bria, F. et al., Protect, prepare and transform Europe – Recovery and resilience post COVID-19, Publications Office of the European Union, https://data.europa.eu/doi/10.2777/47540

European Commission (2021a). Directorate-General for Budget. The EU's 2021-2027 long-term budget and NextGenerationEU – Facts and figures, Publications Office of the European Union, https://data.europa.eu/doi/10.2761/808559

European Commission (2021b). Analysis of the recovery and resilience plan of Portugal. SWD(2021) 146 final. https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52021SC0146

European Commission (2023). Recovery and Resilience Facility Annual Report 2023. COM(2023) 545 final/2. https://commission.europa.eu/system/files/2023-09/COM_2023_545_1_EN_0.pdf

Foray, D., David, P. A., and Hall, B. (2009). Smart Specialisation – The concept, Knowledge for Growth Expert Group Policy Brief No. 9, June. European Commission.

Foray, D.; Goddard, J.; Goenaga Beldarrain, X.; Landabaso, M.; McCann, P.; Morgan, K.; Nauwelaers, C. and Ortega-Argilés, R. (2012). Guide to Research and Innovation Strategies for Smart Specialisation, Luxembourg: Publications Office of the European Union.

Foray, D., Eichler, M., & Keller, M. (2021). Smart specialization strategies—Insights gained from a unique European policy experiment on innovation and industrial policy design. *Review of Evolutionary Political Economy*, 2(1), 83–103.

Griessen, T., & Braun, D. (2008). The political coordination of knowledge and innovation policies in Switzerland. *Science and Public Policy*, 35(4), 277–288. https://doi.org/10.3152/030234208X310338

Grimmer, J., & Stewart, B. M. (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis*, 21(3), 267–297. https://doi.org/10.1093/pan/mps028

Hammitt, J. K., & Zhang, Y. (2013). Combining experts' judgments: Comparison of algorithmic methods using synthetic data. *Risk Analysis: An International Journal*, 33(1), 109-120.

Huang, Y., Li, Y., Mao, J., Li, R., & Zhang, L. (2023). The dynamics of policy coordination: The case of China's science and technology policy-making. *Science and Public Policy*, 50(2), 177–193. https://doi.org/10.1093/scipol/scac058

Humphreys, A., & Wang, R. J. H. (2018). Automated text analysis for consumer research. *Journal of Consumer Research*, 44(6), 1274-1306.

Kelchtermans, Stijn. (2022). Stata script using Latent Semantic Analysis (LSA) to calculate the similarity between Resilience and Recovery Plan measures and national/regional smart specialisation priorities for Portugal (Version V5). Zenodo. https://doi.org/10.5281/zenodo.6560197

Kivimaa, P., Lukkarinen, J., & Lazarevic, D. (2023). Analysis of COVID-19 recovery and resilience policy in Finland: A transformative policy mix approach. *Science and Public Policy*, 50(4), 681–694. https://doi.org/10.1093/scipol/scad016

Krippendorff, K. (1980). *Content analysis: An introduction to its methodology.* Beverly Hills, CA: SAGE Publications.

Landauer, T. K., & Kintsch, W. (2003). *Latent semantic analysis. Automated Essay Scoring: A Cross-disciplinary Perspective*, Routledge.

Magerman, T., Van Looy, B., & Song, X. (2010). Exploring the feasibility and accuracy of Latent Semantic Analysis based text mining techniques to detect similarity between patent documents and scientific publications. *Scientometrics*, 82(2), 289-306.

Marques Santos, A. (2021). Linking the 'Recovery and Resilience Plan' and Smart Specialisation. The Portuguese Case. JRC Working Papers on Territorial Modelling and Analysis No. 05/2021, European Commission, Seville, JRC 126178.

Milgrom, P., & Roberts, J. (1990). The economics of modern manufacturing: Technology, strategy, and organization. *The American Economic Review*, 511-528.

Park, I., Jeong, Y., Yoon, B., & Mortara, L. (2015). Exploring potential R&D collaboration partners through patent analysis based on bibliographic coupling and latent semantic analysis. *Technology Analysis & Strategic Management*, 27(7), 759-781.

Peters, B. G. (2018). The challenge of policy coordination. *Policy Design and Practice*, 1(1), 1–11. https://doi.org/10.1080/25741292.2018.1437946

Plano de Recuperação e Resiliência (2021). Aviso de Abertura de Concurso. https://recuperarportugal.gov.pt/wp-content/uploads/2021/07/Concurso-Ideias-C5-i01.pdf

República Portuguesa (2021). Plano de Recuperação e Resiliência, 22 April 2021, XXII Governo, Portugal. Published at https://www.portugal.gov.pt/pt/gc22/comunicacao/documento?i=recuperar-portugal-construindo-o-futuro-plano-de-recuperacao-e-resiliencia

Rogge, K. S., & Reichardt, K. (2016). Policy mixes for sustainability transitions: An extended concept and framework for analysis. *Research Policy*, 45(8), 1620–1635. https://doi.org/10.1016/j.respol.2016.04.004

Schwarz, C. (2019). lsemantica: A command for text similarity based on latent semantic analysis. *The Stata Journal*, 19(1), 129-142.

Shen, X., Chen, Z., Backes, M., & Zhang, Y. (2023). In chatgpt we trust? measuring and characterizing the reliability of chatgpt. *arXiv preprint* arXiv:2304.08979.

Topkis, D. M. (1998). *Supermodularity and Complementarity*, Princeton University Press, Princeton, NJ.

Tamtik, M. (2017). Policy coordination challenges in governments' innovation policy—The case of Ontario, Canada. *Science and Public Policy*, 44(3), 417–427. https://doi.org/10.1093/scipol/scw074

Terzi, A., Sherwood, M., & Singh, A. (2023). European industrial policy for the green and digital revolution. *Science and Public Policy*, scad018. https://doi.org/10.1093/scipol/scad018

Tosun, J., & Lang, A. (2017). Policy integration: Mapping the different concepts. *Policy Studies*, 38(6), 553–570. https://doi.org/10.1080/01442872.2017.1339239

Trein, P., Meyer, I., & Maggetti, M. (2019). The Integration and Coordination of Public Policies: A Systematic Comparative Review. *Journal of Comparative Policy Analysis: Research and Practice*, 21(4), 332–349. https://doi.org/10.1080/13876988.2018.1496667

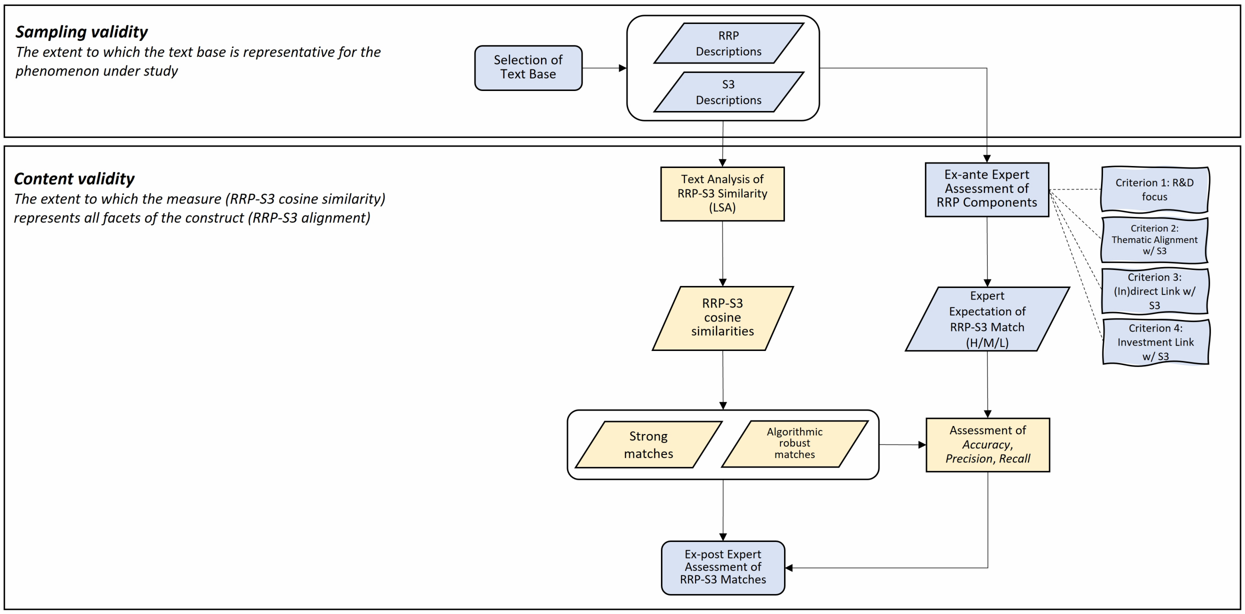
Trein, P., Biesbroek, R., Bolognesi, T., Cejudo, G. M., Duffy, R., Hustedt, T., & Meyer, I. (2021). Policy Coordination and Integration: A Research Agenda. *Public Administration Review*, 81(5), 973–977. https://doi.org/10.1111/puar.13180

Valdez, D., Pickett, A. C., & Goodson, P. (2018). Topic modeling: latent semantic analysis for the social sciences. *Social Science Quarterly*, 99(5), 1665-1679.

Vrana, S. R., Bono, R. S., Konig, A., & Scalzo, G. C. (2019). Assessing the coherence of narratives of traumatic events with latent semantic analysis. *Psychological Trauma: Theory, Research, Practice, and Policy*, 11(5), 521.

Tables and Figures

Figure 1: Schematic Representation of the Integrated Workflow



Note: *Rectangles* indicate process steps; parallelograms indicate inputs and outputs; *oval shapes* indicate entry and end points of the procedure. *Yellow elements* indicate the use of automated text analysis; *blue elements* indicate expert input.

Table 1: Recovery and Resilience Plan (RRP) and Smart Specialisation Strategies (S3) of Portugal

RRP *(20 components, average length: 404 words)*	
Health	Decarbonisation of industry
Housing	Sustainable bioeconomy
Social response	Energy efficiency in buildings
Culture	Hydrogen and renewables
Capitalization and Innovation	Sustainable mobility
Qualifications and skills	Business 4.0
Infrastructure	Quality and sustainable public finance
Forest	Economic justice & business environment
Water management	Public administration digitalization
Sea	Digital school
National and Regional S3 (2021-2027)	
National *(7 priorities, average length: 89 words)*	***Lisboa*** *(8 priorities, average length: 847 words)*
Economy 4.0 and Digital KET	Agrofood
Materials and advanced production technologies	Blue economy
Circular Economy, Energy Transition and Decarbonisation	Creative and cultural industries
Health, biotechnology and food	Mobility and transport
Territory, creativity and brands	Health
Large natural & emerging assets: sea	Tourism and hospitality
Large natural & emerging assets: space tech & earth observ.	Digital Transition
	Transversal Higher Education
Centro *(6 priorities, average length: 189 words)*	***Algarve*** *(6 priorities, average length: 123 words)*
Natural resources and bioeconomy	Tourism and leisure
Material, tooling and production technologies	Sea, fisheries and aquaculture
Digital technology and space	Health, Wellbeing, Life Sciences
Energy and climate	Agrifood, agro-transformation, forest & green biotech
Health and well-being	ICT and cultural and creative industries
Culture, creativity and tourism	Renewable energy
Norte *(8 priorities, average length: 100 words)*	***Alentejo*** *(8 priorities, average length: 304 words)*
Creativity, Fashion and Habitats	Digitalization of economy
Industrialization and Advanced Manufacturing Systems	Circular economy
Agro-Environmental Systems and Food	Sustainable bioeconomy
Sustainable Mobility and Energy Transition	Sustainable Energy
Life Sciences and Health	Mobility & logistics
Territorial Assets and Tourism Services	Tourism services
Resources and Sea Economy	Creative and cultural industries
Technologies, State, Economy and Society	Social innovation and citizenship
Madeira *(6 priorities, average length: 489 words)*	***Azores*** *(4 priorities, average length: 699 words)*
Health and wellbeing	Agriculture and Agroindustry
Agriculture, food and bioeconomy	Sea and blue growth
Circular economy, energy transition, climate change &	Tourism and Heritage

biodiversity	
Digital technologies and economy 4.0	Space and data science
Marine resources and technologies	
Tourism	

Sources: Portuguese RRP and Marques Santos (2021).

The "average length" refers to the section of the policy document that describes the corresponding RRP component or S3 priority.

Table 2: Summary of results for the semantic similarity of the RRP with all S3 policies in Portugal

Smart specialisation strategy (S3)	**Alignment of the RRP with the S3 strategy**[1]	**Reliability of matches**[2]	**Classification performance**[3]					
			Based on 4 expert criteria[4]			*Based on 1 expert criterion (thematic alignment only)*		
			Accuracy	*Precision*	*Recall*	*Accuracy*	*Precision*	*Recall*
National	75%	93%	69%	50%	100%	92%	91%	100%
Lisbon	60%	100%	47%	22%	67%	80%	71%	83%
Norte	60%	100%	44%	22%	50%	100%	100%	100%
Centro	60%	100%	47%	22%	67%	89%	100%	85%
Algarve	45%	89%	60%	29%	67%	88%	86%	86%
Alentejo	55%	100%	50%	14%	50%	100%	100%	100%
Azores	30%	100%	79%	33%	100%	95%	80%	100%
Madeira	70%	100%	37%	8%	100%	88%	82%	100%
Average	57%	98%	54%	25%	75%	92%	89%	94%

[1] The share of RRP measures that is matched with an S3 priority i.e. for which the RRP-S3 cosine similarity for the best matches is higher than the threshold.

[2] The share of RRP-S3 matches for which the algorithm produces consistent results under different parameterizations. Only RRP-S3 matches for which the cosine similarity is above the threshold are included.

[3] The classification performance compares the match based on the algorithm with the expert expectation of a match and results in one of the following outcomes: TP=True Positive; FP=False Positive; TN=True Negative; FN=False Negative. If the expert did not have a clear expectation (high/low) about finding a similar S3 priority for a given RRP component, it is not classified and not taken into account. *Accuracy* = (TP+TN)/(TP+FP+FN+TN); *Precision* = TP/(TP+FP); *Recall* = TP/(TP+FN).

[4] The expert assesses whether an RRP component is aligned with a given S3 priority based on their 1) R&D focus; 2) theme; 3) (in)direct link; 4) investment beneficiaries.

Table 3: Alignment of the national and regional S3 strategies (top-down)

	Alignment with regional S3 strategies		
National S3 priority	**Nr of Regions[1]**	**% of regions**	**Cosine Similarity[2]**
Economy 4.0 AND Digital KET	5	71%	0.69
Advanced Production Materials and Technologies	3	43%	0.78
Circular Economy, Energy Transition and Decarbonisation	5	71%	0.68
Health, Biotechnology and Food	6	86%	0.72
Territory, Creativity and Brands	7	100%	0.78
Large Natural and Emerging Assets - Sea	5	71%	0.69
Major Natural and Emerging Assets - Space and Earth Observation Technologies	4	57%	0.74
Average across national S3 priorities	**5**	**71%**	**0.73**

[1] The number of regions that has at least one regional S3 priority that matches with the national S3 priority. A match requires that the best matching regional S3 priority has a cosine similarity above the matching threshold.

[2] The average cosine similarity of the regional S3 priorities and the national S3 priority, conditional on a match.

Table 4: Alignment of the regional and national S3 strategies (bottom-up)

Regional S3 priority	Best matching national S3 priority	Cosine Similarity
Alentejo		
DIGITALISATION OF THE ECONOMY	Economy 4.0 AND Digital KET	0.59
CIRCULARITY OF THE ECONOMY	Circular Economy, Energy Transition and Decarbonisation	0.61
SUSTAINABLE BIO-ECONOMY	Large Natural and Emerging Assets - Sea	0.53
SUSTAINABLE ENERGY	Circular Economy, Energy Transition and Decarbonisation	0.58
MOBILITY AND LOGISTICS	Major Natural and Emerging Assets - Space and Earth Observation Technologies	0.59
TOURISM AND HOSPITALITY SERVICES	Territory, Creativity and Brands	0.69
CULTURAL AND CREATIVE ECOSYSTEMS	Territory, Creativity and Brands	0.80
SOCIAL INNOVATION AND CITIZENSHIP	Health, Biotechnology and Food	0.50
Alignment with national S3 strategy (share of S3 priorities with match)[1]	*100%*	
Alignment with national S3 strategy (average cosine, conditional on match)[2]	*0.61*	
Algarve		
Tourism and Leisure	Territory, Creativity and Brands	0.64
Sea, Fisheries and Aquaculture	Large Natural and Emerging Assets - Sea	0.57
Agri-food, Agro-Processing, Forest and Green Biotechnology	no match	-
ICT and Cultural and Creative Industries	Territory, Creativity and Brands	0.71
Renewable Energy	no match	-
Health, Well-being and Life Sciences	Health, Biotechnology and Food	0.65
Alignment with national S3 strategy (share of S3 priorities with match)[1]	*67%*	
Alignment with national S3 strategy (average cosine, conditional on match)[2]	*0.64*	
Azores		
Agriculture and agri-business	no match	-
Sea and blue growth	Large Natural and Emerging Assets - Sea	0.50
Tourism and Heritage	Territory, Creativity and Brands	0.65
Space and data science	Major Natural and Emerging Assets - Space and Earth Observation Technologies	0.70
Alignment with national S3 strategy (share of S3 priorities with match)[1]	*75%*	
Alignment with national S3 strategy (average cosine, conditional on match)[2]	*0.62*	
Centro		
Natural Resources and Bioeconomy (Water, Forest, Agro-Food)	Circular Economy, Energy Transition and Decarbonisation	0.65
Materials, Tooling and Production Technologies	Advanced Production Materials and Technologies	0.88
Digital technologies and space	Major Natural and Emerging Assets - Space and Earth Observation Technologies	0.86
Energy and Climate	Circular Economy, Energy Transition and Decarbonisation	0.80
Health and Well-being	Health, Biotechnology and Food	0.87
Culture, Creativity and Tourism	Territory, Creativity and Brands	0.89
Alignment with national S3 strategy (share of S3 priorities with match)[1]	*100%*	
Alignment with national S3 strategy (average cosine, conditional on match)[2]	*0.83*	

Table 4 (continued)

Lisbon		
Agro-food	Health, Biotechnology and Food	0.61
Blue Economy	no match	-
Creative and Cultural Industries	Territory, Creativity and Brands	0.68
Mobility and Transport	Economy 4.0 AND Digital KET	0.57
Health	Health, Biotechnology and Food	0.60
Tourism and hospitality	no match	-
Digital Transition	Economy 4.0 AND Digital KET	0.65
Transversal Higher Education	Economy 4.0 AND Digital KET	0.54
Alignment with national S3 strategy (share of S3 priorities with match)[1]	*75%*	
Alignment with national S3 strategy (average cosine, conditional on match)[2]	*0.61*	
Madeira		
Tourism	Territory, Creativity and Brands	0.81
Marine Resources and Technologies	Large Natural and Emerging Assets - Sea	0.90
Digital Technologies and Economy 4.0	Economy 4.0 AND Digital KET	0.97
Circular Economy, Energy Transition, Climate Action and Biodiversity	Circular Economy, Energy Transition and Decarbonisation	0.88
Agriculture, Food and Bioeconomy	Health, Biotechnology and Food	0.77
Health and Well-being	Health, Biotechnology and Food	0.90
Alignment with national S3 strategy (share of S3 priorities with match)[1]	*100%*	
Alignment with national S3 strategy (average cosine, conditional on match)[2]	*0.87*	
Norte		
Creativity, Mode and Habitats	Territory, Creativity and Brands	0.89
Industrialisation and Advanced Manufacturing Systems	Advanced Production Materials and Technologies	0.96
Agro-environmental systems and food	Circular Economy, Energy Transition and Decarbonisation	0.59
Sustainable Mobility and Energy Transition	Major Natural and Emerging Assets - Space and Earth Observation Technologies	0.65
Life Sciences and Health	Health, Biotechnology and Food	0.81
Territorial assets and Tourism Services	Territory, Creativity and Brands	0.91
Resources and the Economy of the Sea	Large Natural and Emerging Assets - Sea	0.93
Technologies, state, economy and society	Major Natural and Emerging Assets - Space and Earth Observation Technologies	0.81
Alignment with national S3 strategy (share of S3 priorities with match)[1]	*100%*	
Alignment with national S3 strategy (average cosine, conditional on match)[2]	*0.82*	
Average across regions		
Alignment with national S3 strategy (share of S3 priorities with match)[1]	*88%*	
Alignment with national S3 strategy (average cosine, conditional on match)[2]	*0.71*	

[1] The share of S3 priorities of the region that matches with a national S3 priority. A match requires that the best matching national S3 priority has a cosine similarity above the matching threshold (0.49).
[2] The average cosine similarity of the S3 priorities of the region with the national S3 priorities, conditional on a match.

Appendix 1: Pairwise Similarity of Policy Documents (LSA)

The key feature of LSA is that it can judge documents to be similar even if they contain different words, provided the words appear together in a similar context. The approach accomplishes this by, first, determining the frequency by which each word in the vocabulary (i.e., the collection of all words in all documents) appears in a document, resulting in a so-called *document-term matrix* or “bag-of-words” representation of the textual data. These frequencies are then re-weighed – using the *term-frequency-inverse-document-frequency* (tf-idf) – to reduce the weight of words that appear in many documents and thus have little distinctive value (Schwarz, 2019). Finally, the document-term matrix is subjected to a Singular Value Decomposition, which reduces the dimensionality of the data to a set of components[21] that captures most of the variation between the documents. The result is a *document-component matrix* of dimensions , with denoting the documents and the components.[22] The similarity of documents is measured by the cosine similarity of two component vectors. More specifically, the cosine similarity of two documents *d1* and *d2* and their respective document-component vectors δ_{d1} and δ_{d2} is defined as:

Before the LSA analysis can be run, the following issues with respect to the pre-processing of the input documents need to be addressed.

Language. The RRP and S3 text strings should be in the same language to allow for a semantic analysis. Therefore, the descriptions of the RRP components and S3 priorities were machine-translated to English before using them in the analysis.[23]

The descriptions of the RRP components and S3 priorities. There is no universal rule to define the text corpus, but the sample of text should provide a valid description of the phenomenon under study. Krippendorff (1980) defines *sampling validity* as *“the degree to which a collection of data are either statistically representative of a given universe or in some specific respect similar to another sample from the same universe so that the sample can be analysed in place of the universe of interest. In content analysis, it is the degree to which the collection of data contains, with a minimum of bias, a maximum of relevant information about the universe, correcting particularly for the bias in their selective availability”*. In other words, the body of text on which the analysis is performed has to be informative on the focal constructs. While scholars have successfully compared the similarity of scientific papers using only the titles (Schwarz, 2019) and others used entire publications or patent documents (e.g. Magerman, 2010), we use descriptions consisting of several hundred words, taken from the official public documents describing these policies. We comment in more detail on the policy documents, and discuss some examples, in section 4.1.[24]

Cleaning of the documents. Cleaning of the text before running the semantic analysis is needed to avoid that capitalization, punctuation or ‘stopwords’ (“a”, “the”, “be”…) affect the results.[25] For example, without cleaning,

---

[21] Note that the RRP uses the term “components” to refer to specific policy initiatives, which is not to be confused with the components used for the Singular Value Decomposition of the LSA algorithm. In the remainder of the paper, it will be clear from the context what the term refers to.

[22] There is no deterministic rule for choosing the optimal number of components C. As the size of the vocabulary is relatively small – the length of the documents describing RRP measures and S3 priorities is typically only a few hundred words (see Table 1) – we set *C=10* and check the robustness of the findings when doubling it (*C=20*).

[23] We performed a manual inspection of the quality of the machine translation and found it to be high. We also manually adjusted inserted the English translation of some words that remained untranslated and found the impact on the results to be very minor. Finally, note that the method is not sensitive to grammatical mistakes introduced by machine translation since the method is robust against minor changes in the order of words.

[24] Note that the reliance on the official descriptions implies that we cannot extend the descriptions, i.e. we did not alter the official descriptions in any way.

[25] Note that the impact of any remaining stopwords is minimized through the *tf-idf* reweighting.

occurrences of *innovation* and *innovation"* (with an attached closing quote) are regarded as different words.

After the pre-processing of the documents, the cosine similarity between the documents describing the RRP components and the S3 priorities is calculated, where we consider each the of national and 7 regional smart specialization strategies separately.[26]

[26] The analysis uses the LSA implementation in Stata (*lsemantica*) by Schwarz (2019); the analysis script is available in the Zenodo repository (Kelchtermans, 2022).

Appendix 2: Detailed Results

Table 5: Illustration of the national and Lisbon region smart specialization priorities

S3 priority	Word count	Example extract from priority description
National priorities (7)		
Economy 4.0 AND Digital KET	122	*Accelerating the digital transformation of companies in the various user sectors of the national economy. Companies providing technological solutions incorporating digital KETs into the creation of new innovative products and services with strong internationalisation potential.*
Advanced Production Materials and Technologies	88	*Development and use of materials with innovative characteristics (smart, multifunctional, biomaterials, biodegradable and biomimtic materials) and advanced and/or emerging manufacturing technologies (Photonics, Micro and Nano Manufacturing, Industrial Biotechnology, etc.) which, in conjunction with digital technologies, have a wide application in industrial sectors.*
Circular Economy, Energy Transition and Decarbonisation	72	*Exploring opportunities for the development and implementation of innovative solutions contributing to the transition to a competitive and carbon-neutral socio-economic system through a more circular economy promoting material, energy and carbon efficiency, productive efficiency and sustainability of territories.*
Health, Biotechnology and Food	102	*More innovative, robust and resilient, more internationally competitive and higher value-added organisations and value chains, asserting Portugal as an international hub in research and innovation, and as a benchmark for quality and competitiveness in the manufacturing of products and the provision of innovative services in the areas (and their intersections) of health, biotechnology and food.*
Territory, Creativity and Brands	83	*Mobilising intangible dimensions, at the level of symbolic, cultural and creative capital, including intangible dimensions rooted in the territory, and their transformation into value at different stages of value chains, from creation/design to consumption of products and services, with particular relevance in the marketing/promotion phase.*
Large Natural and Emerging Assets - Sea	65	*Mainstreaming decarbonisation, digitalisation and automation processes, with a central role for ports and logistics and the emergence of blue economy hubs. Transition to scale aquaculture with reduced environmental impact.*
Major Natural and Emerging Assets - Space and Earth Observation Technologies	90	*Development of new products (technologies) and services in the construction of rockets and launches, construction of mini, micro and nano-satellites, associated services (management of their operation and the data they provide), enabling earth and sea observation from space and the territory as a major aggregator of application opportunities.*
Priorities for the Lisbon region (8)		
Agro-food	1061	*The Farm to Fork Strategy is aligned with the promotion of the circular economy, with the aim of reducing the environmental impact throughout the food chain and promoting the emergence of new food and feed products. It also relates to stimulating sustainable food consumption and promoting healthy food, being demanding with products imported from outside the EU, fighting food fraud, actions to reduce food waste to help consumers choose healthier and more sustainable diets, and proposals to improve farmers' position in the value chain.*
Blue Economy	565	*Lisbon as a global centre of skills for the blue economy, meeting the interactions between national and regional ocean-related strategies, branding globally recognised marine and maritime products and services, the Portuguese region able to focus and catalyse capacities and resources to assert the role of the ocean in*

		the environmental and energy transitions that will mark the twenty-first century.
Creative and Cultural Industries	977	*Recognising the constraints and fragmentation of the sector, there is huge potential associated with the concentration of equipment (theatres, halls, entertainment rooms, TV studios, etc.), talent (actors, producers, guides, artists, etc.), businesses and educational institutions in the context of their value chain. It highlights the driving role played by the Region, but which could be strengthened, in terms of increasing the public and users of cultural production distributed from Lisbon, as well as the increase in cultural production outside Lisbon.*
Mobility and Transport	756	*The Mobility and Transport domain in Lisbon RIS3 as a system promoting efficiency and organising activities, flows and value chains, creating innovation spaces and opportunities around a global wealth generating project, where the Lisbon Metropolitan Area demonstrates solutions (platforms, services, technology and products) for the world.*
Health	1355	*The proposed package of interventions aims in a structured way to realise the ambition to set up the "Lisbon Health HUB". In particular, the actions will enable the creation of an open science/open data environment, with the provision of specialised services in areas with key local value (social and demographic priorities) but also globally competitive. This will be, inter alia, specifically the case in the area of Ageing: Home support, quality of life programmes, remote monitoring, and integration with post-graduate education programmes (Schools for the World) in the field of health.*
Tourism and hospitality	652	*The ongoing Tagus project (Rede Cais do Tejo project) is an excellent example of this trend and reality. The future of tourism in the Lisbon region requires its compatibility with the communities in which it is located and the creation of a genuine tourist region, which will also serve to reduce tensions with residents of the various municipalities by extending and creating new centres, necessarily including the capacity to innovate and reinvent itself, allowing for new experiences in the same previously visited territory.*
Digital Transition	452	*Public policies in regulation, investments in technological events, among others, should foster the creation of an enabling environment for the digitalisation of the Lisbon Metropolitan Area. One of the examples of strengthening public policies relates to the development of Digital Education.*
Transversal Higher Education	954	*Make the Lisbon Metropolitan Area (AML) a European hub for higher education, research and innovation, leveraging the region's potential to attract, develop and retain talent. This strategic priority is leveraged by public and private commitment to attracting international talent (from students, teachers and researchers), promoting the development of higher education institutions (HEIs) and research and development (R & D) in the region.*

Table 6: Illustration of semantic relationships between words

Dimension	**RRP measure**	**Most frequent word in RRP description**	**Semantically similar words according to the LSA algorithm***
Climate Transition	Decarbonisation of Industry	*energy*	*storage sources carbon low-carbon*
Climate Transition	Energy Efficiency in Buildings	*energy*	*penetration consumption energies retention*
Climate Transition	Hydrogen and Renewable	*energy*	*renewable energies electricity*
Climate Transition	Sea	*blue*	*observatory fish ocean protein microalgae bioremediation*
Climate Transition	Sustainable Bioeconomy	*transition*	*dependence competitiveness industry modernising*
Climate Transition	Sustainable mobility	*transport*	*mobility installed options shipping aeronautical demonstration*
Digital Transition	Business 4.0	*digital*	*technological transformation platforms*
Digital Transition	Digital School	*digital*	*electronic skills educational adoption process promising*
Digital Transition	Economic Justice and the Business Environment	*businesses*	*state administration administrative invest*
Digital Transition	Public Administration — Empowerment, Digitalisation and Interoperability and Cybersecurity	*public*	*challenges responsiveness patterns*
Digital Transition	Quality and Sustainability of Public Finances	*public*	*inequalities aged patterns studies*
Resilience	Capitalisation and Business Innovation	*strategic*	*framework growth standards european national*
Resilience	Culture	*cultural*	*creative intangible heritage artistic audiovisual exhibitions*
Resilience	Forests	*rural*	*control downstream fruit beans locust crop*
Resilience	Housing	*housing*	*households urgent refugees housing*
Resilience	Infrastructure	*traffic*	*organise missing traffic vehicles corridors*
Resilience	Qualifications and Skills	*training*	*digitisation disruptive digitally incubators*
Resilience	Social Responses	*social*	*people personal families family birth reskill*
Resilience	The National Health Service	*health*	*well-being care ageing specialities telemonitoring*
Resilience	Water Management	*water*	*managed river storms utilitydetection floods disaster*

* *Example words that show a high similarity with the most frequent word in the RRP measure description.*

Table 7: Alignment between the RRP and the national S3 strategy of Portugal

	Expert	*Algorithm*		*Expert vs Algorithm*
RRP Component	**Expectation of match**[1]	**Most similar S3 priority**[2]	**Reliable match**[3]	**Classification performance**[4]
1. The National Health Service	high	Health, Biotechnology and Food (0.70, Δ=0.66)	yes	TP
2. Housing	low	- (0.34, Δ=0.26)	-	TN
3. Social Responses	low	Health, Biotechnology and Food (0.62, Δ=0.55)	yes	FP
4. Culture	medium	Territory, Creativity and Brands (0.68, Δ=0.59)	yes	-
5. Capitalisation and Business Innovation	medium	Economy 4.0 AND Digital KET (0.40, Δ=0.12)	yes	-
6. Qualifications and Skills	low	- (0.24, Δ=0.08)	-	TN
7. Infrastructure	low	Territory, Creativity and Brands (0.37, Δ=0.05)	no	FP
8. Forests	medium	Circular Economy, Energy Transition and Decarbonisation (0.47, Δ=0.09)	yes	-
9. Water Management	medium	Circular Economy, Energy Transition and Decarbonisation (0.65, Δ=0.43)	yes	-
10. Sea	medium	Large Natural and Emerging Assets - Sea (0.57, Δ=0.15)	yes	-
11. Decarbonisation of Industry	high	Circular Economy, Energy Transition and Decarbonisation (0.60, Δ=0.45)	yes	TP
12. Sustainable Bioeconomy	high	Circular Economy, Energy Transition and Decarbonisation (0.80, Δ=0.28)	yes	TP
13. Energy Efficiency in Buildings	low	Circular Economy, Energy Transition and Decarbonisation (0.42, Δ=0.25)	yes	FP
14. Hydrogen and Renewable	medium	Circular Economy, Energy Transition and Decarbonisation (0.53, Δ=0.27)	yes	-
15. Sustainable mobility	medium	Circular Economy, Energy Transition and Decarbonisation (0.37, Δ=0.20)	yes	-
16. Business 4.0	high	Economy 4.0 AND Digital KET (0.78, Δ=0.53)	yes	TP
17. Quality and Sustainability of Public Finances	low	- (0.18, Δ=0.05)	-	TN
18. Economic Justice and the Business Environment	low	- (0.27, Δ=0.06)	-	TN
19. Public Administration	low	- (0.33, Δ=0.04)	-	TN
20. Digital School	low	Economy 4.0 AND Digital KET (0.69, Δ=0.48)	yes	FP

[1] Expectation of the expert that the RRP component has a matching S3 priority.

[2] The most similar S3 priority is shown as "-" if none of the RRP-S3 pairwise comparisons yields a cosine similarity above the threshold of 0.35 i.e. the average cosine similarity of all pairs + 1 sd. The cosine similarity of the most similar S3 priority is mentioned in parentheses, joint with the difference Δ between the best and 2nd best match.

[3] The algorithm is considered to yield reliable results if there is overlap in the 2 most similar S3 priorities when using 10 versus 20 components for the singular value decomposition of the LSA algorithm. Reliability is not assessed (-) if none of the RRP-S3 comparisons yields a cosine similarity above the threshold.

[4] The classification performance compares the match based on the algorithm with the expert expectation of a match and results in one of the following outcomes: TP=True Positive; FP=False Positive; TN=True Negative; FN=False Negative, "-" = unclassified (if the expert did not have a clear expectation (high/low) about finding a similar S3 priority).

Table 8: Alignment between the RRP and the S3 strategy for the Lisbon region

	Expert	*Algorithm*		*Expert vs Algorithm*
RRP Component	**Expectation of match**[1]	**Most similar S3 priority**[2]	**Reliable match**[3]	**Classification performance**[4]
1. The National Health Service	low	Health (0.78, Δ=0.35)	yes	FP
2. Housing	low	- (0.66, Δ=0.08)	-	TN
3. Social Responses	low	Health (0.73, Δ=0.30)	yes	FP
4. Culture	medium	Creative and Cultural Industries (0.80, Δ=0.25)	yes	-
5. Capitalisation and Business Innovation	high	Tourism and hospitality (0.82, Δ=0.01)	yes	TP
6. Qualifications and Skills	medium	Health (0.77, Δ=0.06)	yes	-
7. Infrastructure	medium	- (0.53, Δ=0.13)	-	-
8. Forests	low	Agro-food (0.82, Δ=0.07)	yes	FP
9. Water Management	low	Agro-food (0.81, Δ=0.08)	yes	FP
10. Sea	high	Blue Economy (0.81, Δ=0.12)	yes	TP
11. Decarbonisation of Industry	high	- (0.34, Δ=0.04)	-	FN
12. Sustainable Bioeconomy	medium	- (0.60, Δ=0.10)	-	-
13. Energy Efficiency in Buildings	low	- (0.46, Δ=0.02)	-	TN
14. Hydrogen and Renewable	low	- (0.53, Δ=0.03)	-	TN
15. Sustainable mobility	low	- (0.52, Δ=0.02)	-	TN
16. Business 4.0	medium	Digital Transition (0.79, Δ=0.02)	yes	-
17. Quality and Sustainability of Public Finances	low	Health (0.72, Δ=0.03)	yes	FP
18. Economic Justice and the Business Environment	low	- (0.49, Δ=0.00)	-	TN
19. Public Administration	low	Health (0.69, Δ=0.03)	yes	FP
20. Digital School	low	Digital Transition (0.87, Δ=0.08)	yes	FP

[1] Expectation of the expert that the RRP component has a matching S3 priority.

[2] The most similar S3 priority is shown as "-" if none of the RRP-S3 pairwise comparisons yields a cosine similarity above the threshold of 0.67 i.e. the average cosine similarity of all pairs + 1 sd. The cosine similarity of the most similar S3 priority is mentioned in parentheses, joint with the difference Δ between the best and 2nd best match.

[3] The algorithm is considered to yield reliable results if there is overlap in the 2 most similar S3 priorities when using 10 versus 20 components for the singular value decomposition of the LSA algorithm. Reliability is not assessed (-) if none of the RRP-S3 comparisons yields a cosine similarity above the threshold.

[4] The classification performance compares the match based on the algorithm with the expert expectation of a match and results in one of the following outcomes: TP=True Positive; FP=False Positive; TN=True Negative; FN=False Negative, "-" = unclassified (if the expert did not have a clear expectation (high/low) about finding a similar S3 priority).

Appendix 3: Data files and code for analysis

The data files and analysis scripts are publicly available in the Zenodo repository.[27] The included files are the following (BE=Belgium, PT=Portugal):

- ***RRP_S3_docs_PT.xlsx*** and ***RRP_S3_docs_BE.xlsx***: The source documents describing the RRP components and national/regional S3 priorities used as input for the analysis.
- ***match_RRP_S3 PT.do*** and ***match_RRP_S3 BE.do***: Stata (version 15.1) script that reads and cleans the documents, runs the analysis and generates the results.
- ***RRP_S3_results_PT.xlsx*** and ***RRP_S3_results_BE.xlsx***: The main results file with the raw script output as well as tables summarizing the results.

[27] https://doi.org/10.5281/zenodo.6560197