

Towards Future Oriented Collaborative Policy Development for Rural Areas and People

Miloš Ulman¹, Pavel Šimek¹, Jan Masner¹, Pavel Kogut², Tuula Löytty³, Patrick Crehan⁴, Karel Charvát⁵, Antoni Oliva⁶, Stein Runar Bergheim⁷, Milan Kalaš⁸, Denis Kolokol⁸, Tommaso Sabbatini⁸

¹ Department of Information Technologies, Faculty of Economics and Management, Czech University of Life Sciences Prague, Czech Republic

² 21C Consultancy Limited, United Kingdom

³ Smart & Lean Hub Oy, Finland

⁴ Crehan, Kusano & Associates, Belgium

⁵ Czech Center for Science and Society, Czech Republic

⁶ 22SISTEMA, Spain

⁷ Asplan viak internet a.s., Norway

⁸ Kajo s. r.o., Slovakia

Abstract

Rural areas in Europe are at risk due to depopulation, failing generation renewal, and a multitude of influences ranging from market-based, regulatory, to societal and climate changes. As a result, current rural policy is no longer keeping pace with these changes. We propose an advanced rural policy development framework in order to deliver more accurate foresight for rural regions, contributing to new and enhanced policy interventions. The proposed framework combines new quantitative and qualitative epistemological approaches, previously unused unstructured data with traditional research information, grassroots perspective with expert knowledge, current situation analysis with forward looking activities. We argue that by using the proposed methods, policy teams will be able to enhance the effectiveness of their policy making processes, while rural stakeholders will be given the opportunity to become valuable policy influencers and solution co-creators. The ability to quickly experiment and understand the impact of a variety of policy solutions will result in saved time and costs. The framework is part of an ongoing experimental verification and testing in twelve pilot regions across Europe and Israel.

Keywords

Rural areas, policy, European Union, foresight, text mining, system dynamics modelling.

Ulman, M., Šimek, P., Masner, J., Kogut, P., Löytty, T., Crehan, P., Charvát, K., Oliva, A., Bergheim, S. R., Kalaš, M., Kolokol, D. and Sabbatini, T. (2020) "Towards Future Oriented Collaborative Policy Development for Rural Areas and People", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 1, pp. 111-124. ISSN 1804-1930. DOI 10.7160/aol.2020.120110.

Introduction

Rural areas are under pressure. Over the past twenty-five years, rural regions have experienced a rapidly shrinking population as people, especially young adults, have migrated to cities (Westhoek, van den Berg and Bakkes, 2006; Wiebe et al., 2018). The impact of this demographic shift is profound. European Commission reports that only 5.6% of all European farms are run by people younger than 35 while more than 31% of all farmers are older than 65 (DG Agri, 2017). This imbalance creates difficulties for generational renewal and raises concerns about the loss

of valuable skills and knowledge as older, more experienced workers leave the sector. Skills needed to operate a farm have rapidly shifted to highly intensive brain work and business acumen. Additionally, market-based, regulatory and social changes have a strong bearing on how agricultural producers conduct their business. Climate change and environmental degradation increase farmers' responsibility for conservation of natural resources (Van Herzele et al., 2013). In addition to the down hill of agriculture, also the other earning options suffer from the population shift from rural to urban areas e.g. tourism, well-being, bio-based industry. The implication is that

current rural policy is no longer keeping pace with the changing world as well as it used to.

Decision makers have the ability to steer change and in so doing reduce the negative impact thereof, however this requires advanced knowledge of how a particular action, or inaction, will affect people and places, at present and in the future. Generating such knowledge/insight is easier said than done. First, it is difficult to obtain if parties work in isolation from each other; the outcome is best if the effort is a collective one. Despite a growing recognition that societal goals are best achieved when multiple actors join in the policymaking process, a truly participatory setup remains more an exception than the norm (Bourgeois et al., 2017). Second, the required knowledge must shed light on the current state of affairs as well as on what is yet to come, providing a 360-degree view that is rarely present in today's policymaking (Wiebe et al., 2018). This means having a good understanding of what different rural stakeholders want and need; whether measures aimed at addressing these needs are adequate; how the present situation may evolve under different circumstances in the coming decades; what driving forces will be most influential and why; and how will all this affect planet, people, profits and land-use?

The use of text mining is only starting to gain traction in other domains while in rural policy making it is virtually non-existent (Kayser and Blind, 2017). The quantitative-qualitative nexus in futures exploration, despite some signs of strengthening in recent years, continues to be marked by a strong polarisation when it comes to methodological choices. In futures research, a long-standing divide remains in place between strictly quantitative and more qualitative approaches (Fontela, 2000). Outlook studies that report on future developments in specific sectors tend to be based on either one or the other, with little common ground in between. Whilst a hybrid approach has been tested in the past (Greiner et al., 2014; Fortes et al., 2015), it wouldn't be an overstatement to say that a practice of combining the two has yet to hit the mainstream.

This paper aims to propose an advanced rural policy development framework in order to deliver more accurate foresight for rural regions, contributing to new and enhanced policy interventions to improve rural attractiveness as a place to live and work for newcomers and current rural inhabitants. Further, we formulate following research questions. The first question is how to design a participatory policy co-creation process that is able to combine inputs

from experts, policymakers and all relevant rural stakeholders. The second question is how text mining of publicly accessible data can assist with evaluation of the current state of rural development and contribute to modelling future scenarios. Third question is how to integrate both qualitative and quantitative research approaches in foresight in order to overcome its limitations.

The paper is organized as follows. First, we present the PoliRural framework and its baseline in the Materials and methods section. This is followed by a detailed description and discussion of five components of the framework in the Results and discussion section. A summary of paper contributions and future research are presented in the Conclusion.

Materials and methods

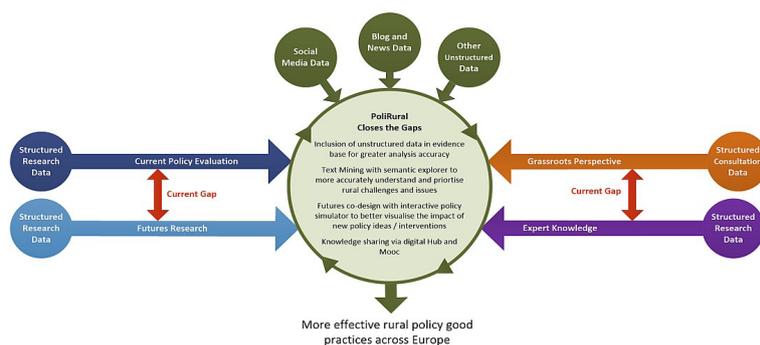
This paper reports on rural policy making approaches and use of text mining. We used desk research to collect data on issues and challenges of rural policy making with a particular attention paid to the current state in the European Union. This served as an input for the design of a new framework called PoliRural. The proposed framework combines new quantitative and qualitative epistemological approaches, previously unused unstructured data with traditional research information, grassroot perspective with expert knowledge, current situation analysis with forward looking activities. The software-based components have been developed and tested as proof-of-concept.

PoliRural framework aims to create innovative and complex rural policy simulators that can be applied by any region to gain insights on the effectiveness of existing measures, as well as the potential impact of new co-designed interventions in a broader context of factors that affect rural places and people. The visualisation of the interplay between rural policy making, futures research and text mining is presented at Figure 1. This will serve as a baseline for creation of an advanced rural policy development framework and tools.

Results and discussion

In this section, we present the proposed PoliRural framework consisting of five major components:

- participatory foresight for rural policy making,
- text mining enabled policy evaluation,
- model-assisted scenario planning process,
- mission oriented transformation,
- PoliRural innovation hub.



Source: own processing

Figure 1: PoliRural framework baseline.



Source: own processing

Figure 2: Components of the PoliRural framework.

The ultimate goal of the PoliRural framework is to increase the attractiveness of rural areas for living, investment, job creation, employment, recreation, cultural heritage (Figure 2).

Assumptions and interdisciplinary considerations

There has been a growing recognition that the practice of foresight can be more inclusive and that citizens can play a bigger role in re-imagining the future. Futures is a growing discipline, with practitioners building expertise by undergoing academic training, participating in conferences and acquiring membership in professional bodies. Despite this strengthening professionalism, futurists themselves are concerned that the lack of diversity in their ranks leads them to envision disproportionately optimistic or pessimistic futures (Nikolova, 2014).

The recognition of the limits of expert foresight is growing alongside efforts by governments, including the EU, to harness the collective capacity of a society to create greater public value (Staman et al., 2017). Furthermore, engagement of a wider range of stakeholders including citizens in foresight exercises is even more justified in the context of policy initiatives aimed at systemic transformation for solving societal problems. Citizens are indeed considered as holders of knowledge needed to understand ‘wicked’ (i.e. complex

and multifaceted) problems. Their engagement is increasingly seen as crucial for ensuring the wider acceptance, adoption and diffusion of solutions purporting to address societal needs (Bourgeois et al., 2017). In the current context of rising populism and the growing perception of democratic deficit at different levels of government, citizen participation in the policymaking process in general and foresight in particular can lead to policies precisely the kind of legitimacy they need.

Another source of potential legitimacy is evaluation of data since it can provide a well-performing policy the evidence base needed to justify its implementation. However, providing a complete evidence base is easier said than done, not least because of the data tsunami – a term describing overwhelming data volumes for example in telecommunication networks (Zander and Mähönen, 2013), medicine (Ackerman, 2014) or astronomy (Berriman and Groom, 2011). Typically, evaluations focus on a small subset of existing data, excluding much of what is available online and offline due to resource and access constraints. Policy makers are challenged by the ever-increasing amount of data at their disposal to help orient policy through evaluation of its success or failure. All policy fields, from agriculture to transport, are affected by the data tsunami and text mining offers timely access to important information which would

otherwise be practically impossible to extract manually.

But even those evaluations that draw on all available sources have their limits. Due to their reliance on historic performance, evaluations can only be applied to policies that have been around for some time in order to assess the impact thereof. How can then one assess the impact of a new policy that needs to replace or update the one which was found to be underperforming? The foresight discipline has an answer, although it differs between scenario-oriented practitioners and those who advocate a more quantitative approach. Both approaches have their advantages and limitations.

The scenario technique is a method for systematically studying a system to create consistent scenarios of the future. An extensive and in-depth critical discussion of scenario building techniques is provided in Bourgeois et al (2017). Scenarios can broaden one's view of the various states that a system may take by presenting alternative futures. Threats and opportunities are jointly identified by stakeholders so that strategies can be based upon advanced knowledge of what may happen to the current state of affairs in several years or indeed decades. Popular though it may be, the scenario method is often criticised on two major grounds (Brose et al., 2013). First, the system's feedback structure is not analysed in great detail. Second, once the scenarios are generated, there is generally little or no information on how the envisaged changes will affect the system as a whole. These are precisely the two aspects that are addressed by the rival method i.e., system dynamics.

System dynamics is based on the idea that the behaviour of a given system and its subsystems can be expressed through the continuous interaction between agents. Dynamic systems modelling examines causes and effects over time, accepting that complex interactions and feedback between subsystems do not occur simultaneously and are not observable in the same space (Bryden, 2010). Systems thinking seeks to understand any system by examining the linkages and interactions between elements that comprise the entire system. The prevailing belief is that parts of a system can be better understood in the context of relationships with each other and with other systems, rather than in isolation (Skyttner, 2005). Although system dynamics proponents claim that they offer better insights into the impacts and interrelationships within a system, their critics argue that they adopt too deterministic view of the future, which frankly is too elusive for any model to capture.

Regardless of what foresight approach is taken, the success of a new policy can be still undermined if it is grounded in old thinking, such as that when the effort is excessively focused on particular sectors – as in traditional industrial policy – rather than on problem-specific societal challenges.

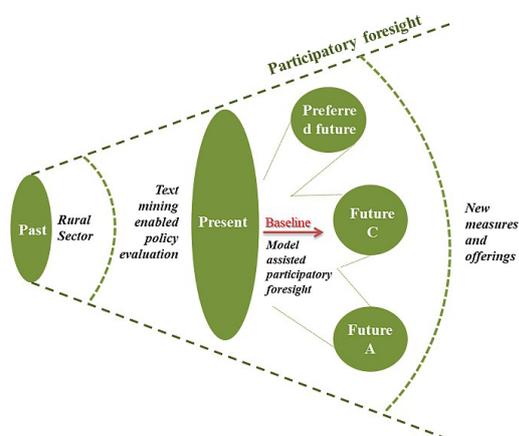
Participatory foresight for rural policy making

We argue that when foresight activity is implemented in a truly participatory way, the result becomes a process of collective learning among participants, leading to a stronger commitment to final results. The PoliRural framework interprets participatory foresight as a powerful combination of strategic anticipatory intelligence, sense-making, visioning, scenario development, systems modelling all coupled with deep participatory engagement that is not limited to the expert community (Bourgeois et al., 2017). As such, foresight is no longer to be treated as a luxury. Indeed, more than ever before, foresight has become an essential prerequisite for proactive, informed and collective actions to stimulate participation of a wider set of actors. For this to happen, the relationship must be underpinned by a shared agenda, an emphasis on value sharing rather than argument, consultative practices based on the principles of inclusion, courtesy and respect.

The main focus of PoliRural foresight is on gaining a well-rounded understanding of change, of how it is happening in the world, how it will play out in the studied regions, and how local/regional policy decisions can influence it for the benefit of grassroot communities. It involves learning an arcane vocabulary relating to macro, meso and micro-trends, trend-breaks and weak signals, drivers and enablers of change and game changers. It is not to be confused with forecasting. It is about understanding not predicting. Tools such as text mining can highlight issues for exploration. Dynamic system modelling can help understand how these issues will evolve. It provides a basis for exploring alternative futures based on scenarios using a participative approach that supports high levels of stakeholder engagement (Hines and Bishop, 2013). Ideal results are obtained when foresight is combined with or embedded in a real and timely local policy process.

The chosen foresight approach will also be strategic and modular. Strategic because it will capture all the required information and facilitate a logical flow between the main stages; modular because at various steps different methods and techniques will be integrated to achieve specific objectives, enabling regional stakeholders to effectively carry

out project activities from start to finish. Inspired by the Framework Foresight method (Hines and Bishop, 2013), we propose the participatory foresight for rural policy making as follows: (1) baseline development from the current situation analysis which culminates in the evaluation of existing policy measures and recommendations for alternative policy options; (2) exploration of future trends and the impact of proposed policy options across space and time, in multiple scenarios and using qualitative and quantitative techniques and tools; (3) implementation by regional stakeholders of selected policy options from the long-list provided using a mission-oriented approach (Figure 3).



Source: adapted from Hines and Bishop (2013).

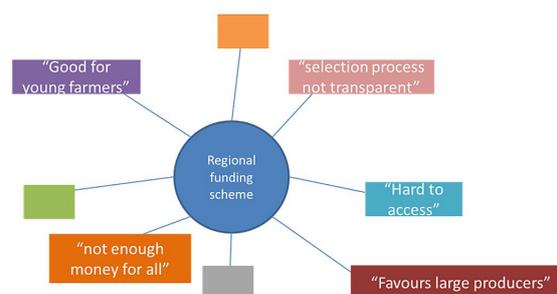
Figure 3: Participatory foresight for rural policy making.

Text mining enabled policy evaluation

We argue that text mining is a technique which is feasible for overcoming the cognition burden that every policy maker faces due to the data tsunami phenomena. Text mining can help to process vast amounts of information from structured and unstructured sources and discover new knowledge at a low cost. Text mining applications can save time on data collection and information processing, allowing decision makers to focus on more important tasks like service delivery. Arguably, the greatest benefit of text mining, when viewed through the prism of participatory foresight (Kayser and Blind, 2017), is that it enhances policy deliberation among citizens, since the application of text-mining techniques to online content found on forums or social media can increase the chances of citizens' voices being heard by decision makers (Chun et al., 2010; Ahn and Bretschneider, 2011).

Text mining will be extensively used at the first foresight stage called current situation analysis. The evaluation will be multi-source and multi-

method, aggregating findings from survey and textual analysis to provide an overview of the current situation that is more complete than the one based on either method alone (Figure 4). The tool itself will be based on heavy-duty knowledge extraction using deep neural networks trained on the large corpus of texts (e.g., EU documents, scientific journals) and adapted to work with regional libraries and languages. Regional libraries are repositories that will contain documents and links to publicly accessible data. The repositories will be curated by researchers in collaboration with policy makers. One of the outputs that this text mining process produces is a semantic tree which can be explored interactively on the PoliRural platform (see PoliRural Innovation Hub section). We will use ANNOY and HDBSCAN for clustering (Melo et al., 2016) and novel Word Mover's Distance for sentence and paragraph similarity analysis (Ye et al., 2016).



Source: own processing

Figure 4: Theoretical text mining.

PoliRural text mining solution has several components. The main part is a set of web crawlers for scraping textual information from online sources using different protocols. Currently, these crawlers collect information from the European Publication Office, Bookshop, EURLEX, CORDIS, DG JRC PUBSY via a number of interfaces (SPARQL, SOAP, OAI, FTP, HTTP) with the help of bespoke harvesters. It is expected that the text mining solution will target only publicly available data sources.

As of today, the harvested repository contains 37 GB of plain text, 6.9 billion tokens and 650,000 unique phrases with cardinality above 20. Tokens are semantic units that result from cutting text into pieces. Depending on the tokenisation strategy used, the results can be quite different e.g., [O'Neill], [Oneil], [neil], [O,neil],[O',neil]. The entire collection passes through the text processing chain based on the open source Python libraries (NLTK, Spacy, Textacy, Tensorflow,

Gensim, Facebook Fasttext). The chain is designed to clean text from artefacts while extracting the relevant metadata i.e. title, year, author, source. The text is then parsed into sentences and phrases before being converted into vectors. Named entities are extracted from the available text using semantic parsing, a multilevel rule-based POS and DEP labelling with entity-type identification. Named entity recognition (NER) locates and classifies named entities found in text into predefined categories such as the names of persons, organizations, locations; expressions of times, quantities, monetary values, percentages, etc (Lin et al., 2019). We will use open source NER algorithms as well as domain specific NER classifiers trained on the PoliRural content repository. Part-of-speech (POS) tagging assigns parts of speech to nouns, verbs, adjectives etc. in order to disambiguate the meaning (Stevenson and Wilks, 2001), while syntactic dependency (DEP) is used to describe the type of syntactic relation that connects the child to the head (Nivre, 2008).

New language models are trained using various Python open-source libraries (e.g., Spacy, Tensorflow, Gensim). This feature is particularly relevant for PoliRural given the fact that it will be employed in a multilingual environment. Newly trained, domain-specific language models can be used for word- or sentence-similarity identification. The text mining tool is also designed to interact with social media. The sources, however, will be limited to platforms that have a public API and offer advanced filtering which is currently Twitter. Data from the streaming API will be stored in the persistent queue which can process large quantities of real-time content. Textual information will then enter the same processing pipelines as described above.

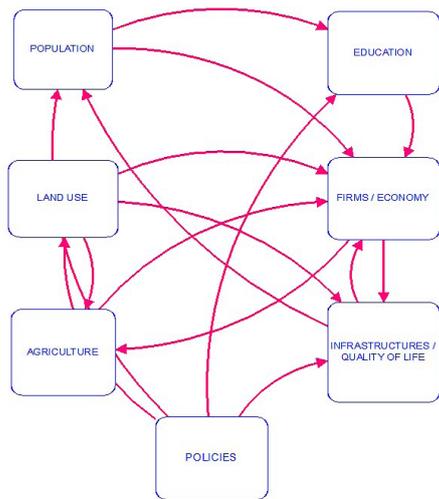
Model-assisted scenario planning process

PoliRural will apply both quantitative and qualitative foresight approaches to create a social learning tool that can help facilitate the discussion among rural stakeholders about the present and particularly future state of agriculture, forestry and other sectors in primary production and beyond. Due to its combined value, the solution is expected to improve decision-making under uncertainty which seems to characterise today's politics, e.g. economic and societal impacts of Covid-19 pandemic, Brexit, etc. To be truly useful for strategic planning, a new innovative solution will be created in a transparent way. It will be developed for rural areas in

cooperation with rural experts and stakeholders, which is an approach promoted by increasing number of studies, e.g., Bryden (2010), Bourgeois et al., (2017), Kano, Fujita and Tsuda (2019). The new solution will capture key facets of rural regions, explain relationships among multiple factors affecting primary production and illustrate potential industry trends and likely impacts of external shocks. Rural stakeholders – government bodies, grassroots community organisations, academia and research institutes – will be able to explore, experiment and visualise what the future may hold in their region. The collaborative process will challenge assumptions, remove prejudices, stimulate debate and improve communication, ultimately helping everyone involved reach a consensus position. This, in turn, would critically inform the formulation of new policies and priorities that can make rural areas more resilient, sustainable and competitive.

One could argue that success and innovation in rural areas is determined by their ability to transform the available capital (human, social, cultural) into new activities and income opportunities. This transformation is sometimes done by farmers themselves, and sometimes by entrepreneurs, or even community organisations not necessarily involved in farming. Policies can be more or less successful in encouraging such transformation, depending in part on the institutional structures and modes of governance at regional levels. Any model that wants to mirror a rural system must therefore consider a whole range of influencing factors and interactions between them.

PoliRural is no exception. Its model will be designed to accommodate a wide range of capitals, policies, demographic, socio-economic and governance mechanisms that might influence the territorial development in different rural and political contexts. The base model (Figure 5) will serve as the conceptual framework for the construction of different regional models. A reappraisal will be carried out when the project starts, but at the time of writing the base model is conceived to include seven modules, each of them adaptable to the local reality of the selected pilot regions. The seven modules are interrelated as shown in Figure 5 (stocks and variables with a discontinuous line indicate they are defined in another module).



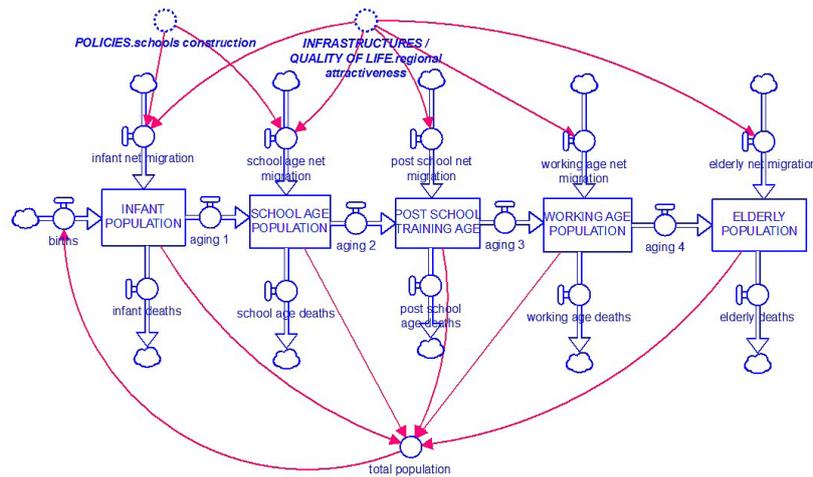
Source: own processing

Figure 5: PoliRural base.

The population module (Figure 6) is based on the aging chain and considers migration in and out in each of the cohorts in which the chain is divided. Depending on data availability, either fertility rate or population growth rate will be used as a source. In the initial base model, attractiveness is considered as having an effect on in/out migration. The precise composition of this variable will be defined according to regional contexts.

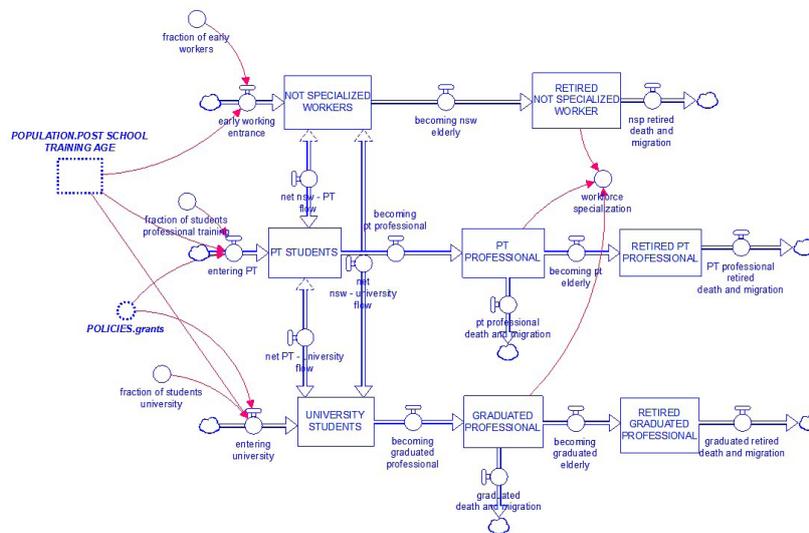
The education module (Figure 7) has a structure of a triple aging chain. Currently, the main output is conceived to be workforce specialisation, which in turn affects the economy module. The model is designed in a way so that different education programs and policies can be easily tested.

The land use module (Figure 8) is concerned with forest land, agricultural land, degraded



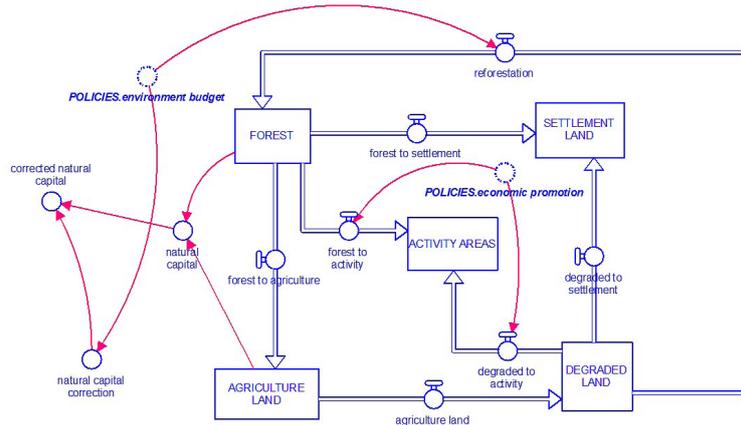
Source: own processing

Figure 6: Population module.



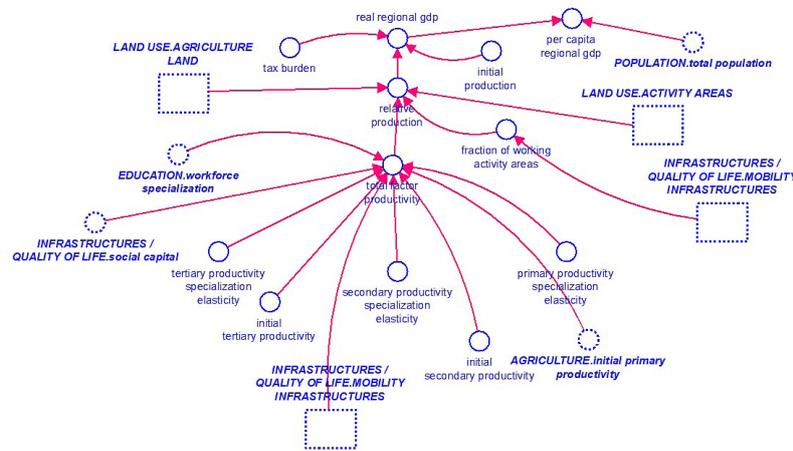
Source: own processing

Figure 7: Education module.



Source: own processing

Figure 8. Land use module.



Source: own processing

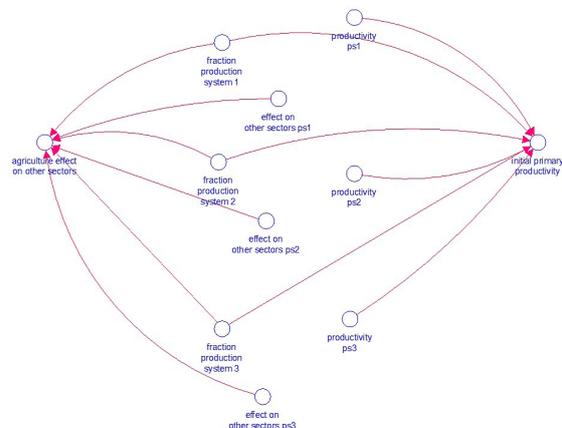
Figure 9. Economy module.

land, settlements land and various activity areas. The main outputs to consider are natural capital and agricultural land, as the unit of agricultural production.

The economy module (Figure 9) is based in a simplified version of the Cobb-Douglas production function. In this version, however, capital is not considered as a production factor. Local perceptions of main problems in the economy and data availability will determine the module's final structure.

The agriculture module (Figure 10) has at its base the production system and basically considers the importance of the sector in rural areas. Each production system has different productivity ratios as well as different effects on the rest of the economy. Three production systems considered in the base model are different management, properties and use of resources. All are dependent on local practices and as such

they differ in social impact, relation with the rest of the economy, profitability, environmental impact, among other things. Examples include intensive livestock farming, extensive farming, cereals, fruit trees, mixed farms.



Source: own processing

Figure 10. Agriculture module.

The quality of life and infrastructures module (Figure 11) considers access to basic services (education, health etc.) and infrastructures such as roads. The infrastructure's stock may be the number of people living within five minutes of a motorway, while quality of life is defined as a function of access to services, the social and the natural capital and the per capita regional gross domestic product. Quality of life is also defining the attractiveness of the area.

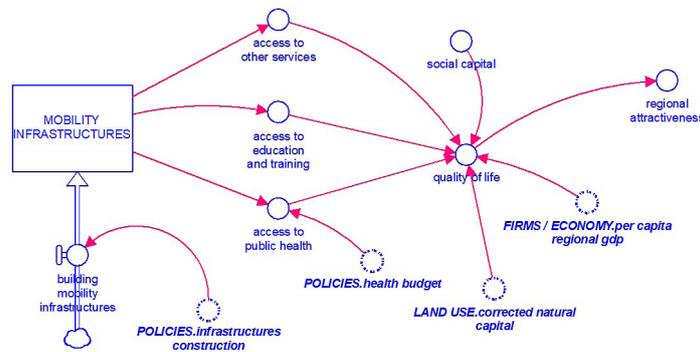
The policies module (Figure 12) is introduced to capture the effects of old and new policies on the different elements within a rural system. The example below considers regional budget and the proportion allocated to the main expenditures. As a consequence of budget allocation, other modules are affected by changes in the policies module. The budget-oriented approach is just one of several options that will be explored in more detail when the project starts.

The model's structure is qualitative to the extent that it reflects local reality, its inherent characteristics and specificities, whereas the statistical input fed into the model adds to its quantitative side. For the model to work as intended, deep domain expertise in system dynamics must interface with grassroots knowledge possessed by regional actors. The model views the present state of affairs as a product of interaction between different agents and sub-systems and tests the impact of proposed

policies on these interactions over time under different scenarios. The three preliminary scenarios to be explored during workshops with regional stakeholders are:

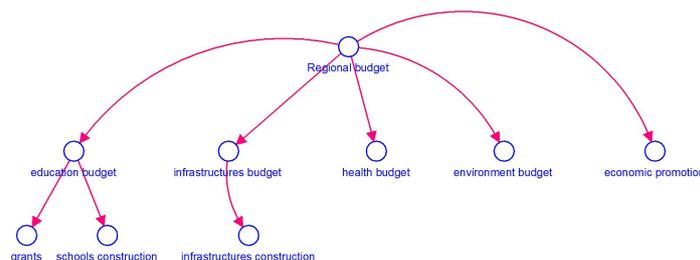
- Business as usual scenario (BAS). BAS shows the evolution of trends observed so far. The trends will reflect the main issues local communities consider to be important
- Local plausible scenarios. This group of scenarios is about the risks and opportunities that the community identifies as plausible in the near future. Scenarios will propose measures to promote opportunities and avoid the risks
- Policy change scenario. The policy change scenario focuses on programs and policies that can help improve local competitiveness and quality of life

In conceptualising and formalising the linkages between policies, farming and land use, production and education, economic and social performance of pilot regions until 2040, PoliRural will initially use Stella Architect which offers a practical set of tools for visualising and communicating how complex systems and ideas work overtime and space. Like other systems models, the core elements of the Stella model are stocks and flows (flows add to or subtract from stocks) and the feedback loop between the two. As the project progresses,



Source: own processing

Figure 11. Quality of life and infrastructure module.



Source: own processing

Figure 12. Policies module.

PoliRural will migrate to Open Source system dynamics software as a base to create a unique innovative software solution for rural policy making.

Mission oriented transformation

Being complex and multifaceted, rural challenges, such as ageing, declining population, unemployment, climate change, droughts, flooding, decreasing biodiversity and environmental degradation require a coordinated and targeted policy response. Traditional approaches that do not deliver in terms of innovation results and solutions should be replaced with a policy approach that can actively contribute to rural change in a direction that is favoured and can be shaped by rural communities. Mission-oriented policy, one of the central concepts of Horizon Europe (European Commission, 2018), offers such an approach. Mission initiatives are characterised by a well-defined goal, direction and timeline, and are typically focused on solving wicked and complex societal challenges and system transformation. Mission-oriented initiatives, being bottom-up, bold and ambitious, as well as cross-sector, cross-discipline and cross-actor oriented, differ in their design, governance and implementation from other types of policies (Fisher et al., 2018; Mazzucato, 2018). Policymakers and stakeholders in PoliRural study areas, supported by the wider PoliRural consortium, will experiment with mission concepts and elaborate mission modalities geared towards their specific rural challenges, needs and context.

PoliRural Innovation Hub

The aim of the PoliRural Digital Innovation Hub (DiH) is to offer a public user interface and introduction to the innovations of the project. To this end, the DiH entry point is built on a content management system. Furthermore, the DiH will provide four distinctive sections, or spaces, that cater for both internal and external users:

1. An interaction space with forums, dialogue and Wiki capabilities to support stakeholder interaction
2. A learning space for Massive Open Online Courses to facilitate dissemination and uptake of knowledge and methodology developed through the project
3. An experimentation space for testing analytics and visualization including text mining and system dynamics based on real data
4. A development and hosting space for creating virtual instances of the shared reference

to be used by each pilot when developing their applications.

Below these high-level functional requirements there is a lot of implicit functionality that in some cases can be supplied by more than one technology component. In order to identify how the DiH can best serve the pilots, the work would need to start off by a detailed analysis of pilots in terms of context, data availability and requirements, analytics requirements and pre-existing tools and technologies that the DiH must be capable of interfacing with.

Relevant EU and international activities

As the PoliRural framework has been designed to be implemented as a research and innovation action project, we have identified a number of initiatives whose results are relevant for the project and as such will be incorporated in whole or in part into the modus operandi when the project starts. These initiatives fall under three main concepts outlined above (participatory foresight, text mining enabled policy evaluation and model assisted participatory scenario building) plus one is related to gender.

Participatory foresight

European Foresight Platform (EFP) is an initiative supported by the European Commission that aims to bring together different communities and individual professionals to share their knowledge about foresight, forecasting and other methods of future studies (EFP, 2010). EFP contains a wealth of information on foresight methodology which can also inform our framework.

Pastoral Properties Futures Simulator (PPFS), a dynamic systems model developed within a participatory action research partnership with the pastoral industry of Australia's Northern Territory (CDU, no date). The process behind PPFS development is well documented in a number of articles (Greiner et al., 2014), and its particularly qualitative elements are accompanied with tried and tested good practices which would be useful for implementation of the PoliRural framework.

Text mining enabled policy evaluation

AGROVOC is a vocabulary developed by FAO of over 35,000 concepts and 671,000 terms in different languages, covering areas such as nutrition, agriculture, fisheries, forestry and environment (AIMS, 2018). PoliRural will adopt AGROVOC as a tool for the initial word embeddings for language models which have not been fully trained i.e. trained on human annotated text only.

FASTPARSE is an ERC funded project that aims to develop fast parsers to improve the analysis and meaning of extracted textual data (Grupolys, 2017). FASTPARSE results will be used to improve PoliRural text mining, semantic explorer, specifically the text processing chain where text will be parsed into sentences and phrases before being converted into vectors.

OpenMinTed is an H2020 that sets out to create an open, service-oriented online infrastructure for text and data mining. OpenMinted results are interesting for PoliRural mainly from the sustainability point of view (ARIS, 2015). OpenMinTed platform will offer a space for PoliRural to publish its text mining tool for a wider uptake by the community. Specifically, PoliRural will be able to publish its content metadata and transfer standards, service metadata and pipelining, IPR and licensing restrictions.

PERCEIVE is an H2020 project that investigates how much citizens living in different parts of Europe feel European, and to what extent this feeling can be attributed to the implementation of the European Cohesion Policy (UNIBO, 2016). The project has developed a rigorous evaluation methodology that contains useful pointers for PoliRural's own evaluation task. Also, the project results will feed directly into the quality of life module that will be developed as part of the system dynamics work stream.

Model-assisted scenario development

TOP-MARD was an FP6 project that used systems dynamics thinking and tools to create a policy model of multifunctional agriculture and rural development (CORDIS, 2013). The developed model provides solid conceptual, scientific and technical cues for building PoliRural own model.

The System Dynamics International Society is an international, non-profit organisation devoted to encouraging the development and use of system dynamics and systems thinking around the world. The society has a Special Interest Group (SIG) on Agriculture & Food that has developed many advanced models covering different aspects of the rural system, from agents to policies to social capital (SDS, 2018).

Conclusion

The paper attempted to make several important contributions. First, we analysed the limits of rural policy development and evaluation

methods and proposed a new approach based on text mining. The text mining tool will feed additional data into the present situation evaluation and future scenarios modelling while relieving researchers from the cognitive burden. Second, inspired by the tried and validated Framework Foresight method, we proposed the participatory foresight for rural policy making approach that combines work with a broad spectrum of stakeholders on policy evaluation and needs analysis, text mining and system dynamics modelling. Third, we argue that future scenario models should be of qualitative nature but, at the same time, fed with quantitative statistical data. This will allow for interaction between domain experts and grassroots stakeholders which should guarantee high precision of the models. All projected contributions will be experimentally tested and verified in a research and innovation action project PoliRural. The project started in June 2019 for a three-year period and was financed under the Horizon 2020 programme.

Despite presenting the contributions based on research in progress, we are able to draw two practical implications for rural stakeholders, national and EU level policy makers. By adopting the proposed PoliRural framework, policy teams will be able to enhance the effectiveness of their policy making processes, while rural stakeholders will be given the opportunity to become valuable policy influencers and solution co-creators. The ability to quickly experiment and understand the impact of a variety of policy solutions will result in saved time and costs. The impact will be demonstrated and extrapolated from twelve pilot sites across Europe and Israel that will use the PoliRural framework and tools for real life policy scenarios, enabling measurement of both quantifiable and qualitative impact measures on the outcomes.

Acknowledgements

This research has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 818496. More details about the project are available at www.polirural.eu.

The results and knowledge included herein have been obtained owing to support from the following institutional grant. Internal grant agency of the Faculty of Economics and Management, Czech University of Life Sciences Prague, grant no. 2019MEZ0004 – “Machine Learning Classification model to Identify User Activities During Usability Testing“.

Corresponding authors

Ing. Milos Ulman, Ph.D.

Department of Information Technologies, Faculty of Economics and Management

Czech University of Life Sciences Prague, Kamycka 129, 165 00 Prague - Suchdol, Czech Republic

Phone: +420 224 382 769, E-mail: ulman@pef.czu.cz

References

- [1] Ackerman, M. J. (2014) "The medical data tsunami", *The Journal of medical practice management: MPM*, p. 406. ISSN 8755-0229.
- [2] Ahn, M. J. and Bretschneider, S. (2011) "Politics of e-government: E-government and the political control of bureaucracy", *Public Administration Review*, Vol. 71, No. 3, pp. 414-424. DOI 10.1111/j.1540-6210.2011.02225.x.
- [3] AIMS (2018) AGROVOC.
- [4] ARIS (2015) "OpenMinted". [Online]. Available: <http://openminted.eu/about/> [Accessed: Dec. 10, 2018].
- [5] Berriman, G. and Groom, S. (2011) "How will astronomy archives survive the data tsunami?", *Communications of the ACM*, pp. 52-56. ISSN 0001-0782. DOI 10.1145/2043174.2043190.
- [6] Bourgeois, R., Penunia, E., Bisht, S. and Boruk, D. (2017) "Foresight for all: Co-elaborative scenario building and empowerment", *Technological Forecasting and Social Change*, Vol. 124 (C), pp. 178-188. ISSN 0040-1625. DOI 10.1016/j.techfore.2017.04.018.
- [7] Brose, A., Gausemeier, P., Vierhaus, I., Fügenschuh, A. and Selinger, G. (2013) "A system dynamic enhancement for the scenario technique", *Proceedings of 11th Global Conference on Sustainable Manufacturing*, pp. 561-566, Universitätsverlag der TU Berlin.
- [8] Bryden, J. M. (2010) "Using system dynamics for holistic rural policy assessments and data envelopment analysis for evaluation of comparative policy efficiency at regional level", *118th Seminar, August 25-27, 2010, Ljubljana, Slovenia*, European Association of Agricultural Economists. DOI 10.22004/ag.econ.94618.
- [9] CDU (no date) Pastoral Properties Futures Simulator.
- [10] Chun, S., Shulman, S., Sandoval, R., Hovy, E. (2010) "Government 2.0: Making connections between citizens, data and government", *Information Polity*, Vol. 15, No. 1, 2, pp. 1-9. ISSN 1570-1255. DOI 10.3233/IP-2010-0205.
- [11] CORDIS (2013) Towards a Policy Model of Multifunctional Agriculture and Rural Development [TOP-MARD].
- [12] DG Agri (2017) Young farmers in the EU – structural and economic characteristics.
- [13] EFP (2010) "European Foresight Platform". [Online]. Available: <http://www.foresight-platform.eu> [Accessed: Dec. 10, 2018].
- [14] European Commission (2018) "Horizon Europe - the next research and innovation framework programme". [Online]. Available at: https://ec.europa.eu/info/horizon-europe-next-research-and-innovation-framework-programme_en [Accessed: Dec. 5, 2018].
- [15] Fisher, R., Chicot, J., Domini, A., Polt, W., Turk, A., Unger, M. and Goetheer, A. (2018) "Mission-oriented research and innovation: Inventory and characterisation of initiatives-Study", European Commission EC. ISBN 978-92-79-85793-5. DOI 10.2777/697082.
- [16] Fontela, E. (2000) "Bridging the gap between scenarios and models", *Foresight*, Vol. 2, No. 1, pp. 10-14. ISSN 1463-6689. DOI 10.1108/14636680010802447.
- [17] Fortes, P., Alvarenga, A., Seixas, J. and Rodrigue, J. (2015) "Long-term energy scenarios: Bridging the gap between socio-economic storylines and energy modeling", *Technological Forecasting and Social Change*, Vol. 91, pp. 161-178. ISSN 0040-1625. DOI 10.1016/j.techfore.2014.02.006.

- [18] Greiner, R., Puig, J., Huchery, C., Collier, N. and Garnett, S. T. (2014) "Scenario modelling to support industry strategic planning and decision making", *Environmental Modelling & Software*, Vol. 55, pp. 120-131. ISSN 1364-8152. DOI 10.1016/j.envsoft.2014.01.011.
- [19] Grupolys (2017) Fast Parse.
- [20] Van Herzele, A., Gobin, A., Van Gossum, P., Acosta, L., Waas, T., Dendoncker, T. and de Frahan, B. H. (2013) "Effort for money? Farmers' rationale for participation in agri-environment measures with different implementation complexity", *Journal of Environmental Management*, Vol. 131, pp. 110-120. ISSN 0301-4797. DOI 10.1016/j.jenvman.2013.09.030.
- [21] Hines, A. and Bishop, P. C. (2013) "Framework foresight: Exploring futures the Houston way," *Futures*, Vol. 51, pp. 31-49. ISSN 0016-3287. DOI 10.1016/j.futures.2013.05.002.
- [22] Kano, E., Fujita, Y. and Tsuda, K. (2019) "A Method of Extracting and Classifying Local Community Problems from Citizen-Report Data using Text Mining", *Procedia Computer Science*, Vol. 159, pp. 1347-1356. ISSN 0016-3287. DOI 10.1016/j.procs.2019.09.305.
- [23] Kayser, V. and Blind, K. (2017) "Extending the knowledge base of foresight: The contribution of text mining", *Technological Forecasting and Social Change*. Vol. 116, pp. 208-215. ISSN 0040-1625. DOI 10.1016/j.techfore.2016.10.017.
- [24] Lin, J. C.-W., Shao, Y., Fournier-Viger, P. and Hamido, F. (2019) "BILU-NEMH: A BILU neural-encoded mention hypergraph for mention extraction", *Information Sciences*, pp. 53-64. ISSN 0040-1625. DOI 10.1016/j.ins.2019.04.059.
- [25] Mazzucato, M. (2018) "*Mission-oriented research & innovation in the European Union*", Brussels: European Commission. ISBN 978-92-79-79918-1.
- [26] Melo, D., Melo, D., Toledo, S., Mourão, F., Sachetto, R., Andrade, G., Ferreira, R., Parthasarathy, S. and Rocha, L. (2016) "Hierarchical Density-Based Clustering Based on GPU Accelerated Data Indexing Strategy", *Procedia Computer Science*, pp. 951-961. ISSN 1877-0509. DOI 10.1016/j.procs.2016.05.389.
- [27] Nikolova, B. (2014) "The rise and promise of participatory foresight", *European Journal of Futures Research*, Vol. 2, No. 1, pp. 33. ISSN 21952248. DOI 10.1007/s40309-013-0033-2.
- [28] Nivre, J. (2008) "Algorithms for Deterministic Incremental Dependency Parsing", *Computational Linguistics*, pp. 513-553. E-ISSN 1530-9312, ISSN 0891-2017. DOI 10.1162/coli.07-056-R1-07-027.
- [29] SDS (2018) "*Special Interest Groups (SIGs)*". [Online]. Available: <https://www.systemdynamics.org/society-sigs> [Accessed: Dec. 10, 2018].
- [30] Skyttner, L. (2005) "*General systems theory: Problems, perspectives, practice*", World scientific. 2nd edition. ISBN 978-981-4479-98-1. p. 536. DOI 10.1142/5871.
- [31] Staman, J., Berloznik, R., Weber, M., Salo, A., and Havas, A. (2017) "*Report of the Expert Group 'Strategic Foresight for R&I Policy in Horizon 2020'*", Background papers, Publications Office of the European Union. ISBN 978-92-79-68543-9. DOI 10.2777/768916.
- [32] Stevenson, M. and Wilks, Y. (2001) "The Interaction of Knowledge Sources in Word Sense Disambiguation", *Computational Linguistics*, pp. 321-349. E-ISSN 1530-9312, ISSN 0891-2017. DOI 10.1162/089120101317066104.
- [33] UNIBO (2016) "*PERCEIVE*". [Online]. Available: <https://www.perceiveproject.eu/about-the-project/> [Accessed: Dec. 10, 2018].
- [34] Westhoek, H. J., van den Berg, M. and Bakkes, J. A. (2006) "Scenario development to explore the future of Europe's rural areas", *Agriculture, Ecosystems and Environment*, pp. 7-20. ISSN 0167-8809. DOI 10.1016/j.agee.2005.11.005.
- [35] Wiebe, K., Zurek, M., Lord, S., Brzezina, N., Gabrielyan, G., Libertini, J., Loch, A., Thapa-Parajuli, R., Vervoort, J. and Westhoek, H. (2018) "Scenario Development and Foresight Analysis: Exploring Options to Inform Choices", *Annual Review of Environment and Resources*, pp. 545-570.

ISSN 1543-4938. DOI 10.1146/annurev-environ-102017-030109.

- [36] Ye, X., Shen, H., Ma, X. and Bunescu, R. C. (2016) “From Word Embeddings to Document Similarities for Improved Information Retrieval in Software Engineering”, *2016 IEEE/ACM 38th International Conference on Software Engineering (ICSE)*, ACM, pp. 404-415. DOI 10.1145/2884781.2884862.
- [37] Zander, J. and Mähönen, P. (2013) “Riding the data tsunami in the cloud: myths and challenges in future wireless access”, *IEEE Communications Magazine*, pp. 145-151. E-ISSN 1558-1896, ISSN 0163-6804. DOI 10.1109/MCOM.2013.6476879.