European
Commission

Business Case

Stakeholder consultation analytics

Date: 16/06/2016
Doc. Version: 1.00
Template Version: 2.5

PM²
PROJECT MANAGEMENT
METHODOLOGY

This template is based on PM² V2.5

For the latest version of this template please visit the PM² Wiki

Document Control Information

Settings	Value
Document Title:	Business Case
Project Title:	Stakeholder consultation analytics
Document Author:	Nikolaos Loutas
	Eva Cobos Cortina
Project Owner:	Angelo Tosetti
Project Manager:	Marco Fichera
Doc. Version:	1.00
Sensitivity:	<Public>
Date:	16/06/2016

Document Approver(s) and Reviewer(s):

NOTE: All Approvers are required. Records of each approver must be maintained. All Reviewers in the list are considered required unless explicitly listed as Optional.

Name	Role	Action	Date
Marco Fichera	Project Manager	*Approved*	16/06/2016

Configuration Management: Document Location

The latest version of this controlled document is stored in

https://webgate.ec.europa.eu/CITnet/confluence/x/0AFWH

TABLE OF CONTENTS

1 CONTEXT

1.1 Situation Description

For almost one decade, the European Commission has been promoting citizen engagement in political decision making via a number of research and development projects co-funded under the FP7, Horizon 2020 and CIP programmes. In the context of EU policy making, the European Commission systematically consults citizens and other stakeholders on major policy initiatives. The European Commission currently runs stakeholder consultations for 12 weeks for initiatives that have an impact assessment or evaluation. Following the adoption of the Better Regulation Guidelines[1], the Commission is planning to consult stakeholders on a more frequent and effective way, and at all stages of the policy-making process.

European Commission DGs manage stakeholders' feedback collection using EUsurvey, but are also using other channels, as DG GROW did publicly requesting feedback on draft legislation via an e-mail form (collecting free text feedback messages), from which one collected over 25000 responses to analyse.

Similar practices are followed also by national, regional and local governments in the EU-28. For example:

- In Greece, the initiative Opengov.gr focus on 3 elements: Open calls for the recruitment of public administration officials, electronic deliberation on draft legislation and policy initiatives by the government which can be commented by citizens, and an open innovation initiative, Labs OpenGov, that brings together ideas and proposals from citizens, the public and the private sectors.

- In Slovenia, the platform predlagam.vladi.si offers the opportunity for registered users to make proposals which are forwarded to the corresponding ministry if they receive a minimum support from other registered users.

- In Estonia, the platform osale.ee is one of the initiatives in which the registered users can make comments on draft legislation and make proposals which are forwarded to the corresponding ministry if they receive a minimum support from other registered users.

- In Bulgaria, the platform strategy.bg publishes all draft legislation prior to the submission to the Parliament and gives the opportunity to citizens to make comments for 14 days, as well as to make proposals on any topic.

- In France, the platforms budgetparticipatif.paris.fr and idees.paris.fr give the opportunity for Paris residents to make proposals to allocate 5% of the yearly budget, and give ideas on specific topics.

- In Spain, the platform decide.madrid.es gives the opportunity to the registered residents to make proposals on ideas to implement on any topic, give their opinion on specific topics and make a proposal on budget allocation.

Depending on the topic discussed, stakeholder consultations attract remarkable interest, which is proven, by the, often overwhelming, number of responses and comments collected. For popular policy topics related to the environment, employment or innovation, several thousands of comments and contributions can be collected. Those comments are, in their vast majority, in free text and can be in any of the 24 official languages of the EU.

Once a consultation is closed, it is the responsibility of the public administrations of the EU policy officers to process and make sense of the collected contributions, and elicit actionable evidence which can be provided as input to the policy makers. Our analysis of numerous existing initiatives (for more details on this, we refer the reader to *D02.01 Participatory knowledge for supporting decision making Study Final Report* reference in *Appendix 1: References and Related Documents*) revealed the following challenges:

- Difficulties in processing manually hundreds, and often thousands, of free-text comments;

- Difficulties in eliciting opinions, detecting trends, removing noise, identifying duplicates and campaigns, defining the sentiment, identifying and categorising similar comments and contributions;

---

[1] http://ec.europa.eu/smart-regulation/guidelines/toc_guide_en.htm

- Difficulties in text analysis because of multilingualism.

These challenges limit the value that can be created from the participation of stakeholders in public consultations, and can put at stake the successful implementation of initiatives such as the Better Regulation Guidelines, which as part of its "Conduct consultation work" phase of the proposed open consultation process is clearly putting emphasis also on analysing the input of stakeholders. As it stands now, the analysis, process and translation of contributions are labour-, time- and expertise-intensive activities. Their automation would have a significant positive impact on the organisation, the effectiveness and the efficacy of stakeholder consultations.

1.2 Situation Impact

The current situation has a negative impact on processes supporting policy making, such as the "Conduct consultation work" phase of the proposed open public consultation process under the Better Regulation Guidelines, as follows:

- **Analysing stakeholders' collected input and feedback manually.** The existing stakeholders' consultation analysis process is mostly manual (i.e. error-prone) and inconsistent across the European Commission DGs or public administrations, involving usually several people. This means that contributions may be left out by accident, campaigns may not be detected, the analysis may be biased by personal opinions or links may not be drawn as different people are looking at different contributions, etc.
- **Time-consuming process.** The current process is not resource-efficient, as each consultation's responses need to be analysed and for open questions the analysis is currently not always automated. According to a study commissioned by the European Commission's Secretariat General on the resources' capacities of consultations' manual analysis, the independent consultancy company Risk & Policy Analysts Limited (RPA) calculated[2] a person can analyse 5 answers per hour when comments are being sampled or relate to different issues, and 10 per hour where the comments relate to the same issue. When considering over 25000 contributions as DG GROW collected on their draft legislation (see the details on the example in section 1.1 Situation description), the time required for the analysis would be of 2500 hours (over 300man-days). There is therefore a major need for the process to be as optimised as possible, which can be achieved through text mining tools.
- **Outreach.** Due to the digital divide the participation of stakeholders on consultations is not evenly distributed across societal groups and geographic regions, thus the responses might not be representative of the population affected by the consulted topic.

The current situation has a negative impact on the stakeholders involved (including residents, citizens and organisations), as follows:

- **Trust.** Existing feedback mechanisms do not provide transparency into how the information collected is analysed, used and considered in policy making (if at all).

---

[2] Workshop on *Analysis of data received in stakeholder consultation activities*, 28/04/2016 at building Berlaymont.

2 EXPECTED OUTCOMES

The expected outcomes derived from the implementation of the proposed solution should aim at fulfilling the challenges identified in section 1.2. More specifically, policy officers should be able to:

- Identify the different opinions expressed in the comments and contributions;

- Detect trends in the comments and contributions;

- Detect spam, duplicate comments and campaigns;

- Be aware of the sentiment (ideally beyond simply positive/neutral/negative) of a particular comment or of all the comments and contributions;

- View together similar or related comments and contributions, e.g. identify arguments and counterarguments;

- Have access to statistics on the comments and contributions;

- Prepare the analysis results in formats easy and flexible enough to be visualised using different ways and tools.

3   POSSIBLE ALTERNATIVES

This section describes known or potentially available alternative solutions to tackle the described situation. From the list of the potential alternatives, one of them is clearly chosen. The chosen alternative is detailed in Section 4.

For each identified alternative, a general description, a SWOT analysis and a qualitative assessment is provided. The SWOT analysis provides the major Strengths, Weaknesses, Opportunities and Threats as perceived by the stakeholders considering the organisational impact, the financial impact, the timing impact and the associated risks.

Three possible alternatives have been considered to tackle the situation and meet the objectives:

- Alternative A: No action
- Alternative B: Use of existing solutions as-is
- Alternative C: Develop text mining solutions for analysing stakeholder consultations.

3.1   Alternative A: No action

In the baseline scenario, policy officers will continue to carry out their work following the current approach explained in section **Error! Reference source not found.**.

SWOT Analysis

Strengths	Weaknesses
• No need for investments	• Policy officers are overloaded with the amount of comments to process. • Policy officers will continue not being able to ensure the completeness of the analysis. • The manual process is error-prone, time consuming and very expensive.
Opportunities	**Threats**
N/A	• Stakeholders will be discouraged from participating in consultations, citizen participation initiatives will be put at risk. • Valuable information is lost and does not reach the policy lifecycle.

For obvious reasons, this alternative does not raise itself as the best option since it will not contribute to the achievement of the expected outcomes.

3.2   Alternative B: Use existing consultation analysis solutions as-is

A second alternative entails using existing solutions, such as DORIS of DG CNECT, that allow the analysis of comments and contributions received in stakeholder consultations.

SWOT Analysis

Strengths	Weaknesses
• Limited additional investments. • Possibility to rely on a solution that proves that can deliver. • In house expertise exists. • The expected outcomes can be partly achieved.	• Existing solutions have their shortcomings in terms of analytic capabilities, visualisation, and multilingualism support. • It is not clear how existing solutions can scale to support more requests and increased demand/use.
Opportunities	**Threats**

• Synergies with DORIS of DG CNECT or other solutions used by Commission services. • Included in the Better Regulation toolbox.	• Using existing solutions at a larger scale would raise questions related to scalability, service governance and support.

This alternative can provide some quick wins with a relatively low investment. However, looking at the long term, it does not raise itself as the best option since existing solutions are not always designed with scalability in mind. Hence, adaptations and further development will be required in order to ensure that existing solutions can meet the growing demand. This takes us to the next alternative.

3.3 Alternative C: Develop text mining solutions for analysing stakeholder consultations

A third alternative entails extending reusing and existing solutions, such as DORIS of DG CNECT, that allow the analysis of comments and contributions received in stakeholder consultations, to create scalable and reusable text mining solutions for analysing stakeholder consultations.

SWOT Analysis

Strengths	Weaknesses

• Improve the analysis of contributions by:
  - Identifying the different opinions expressed in the comments and contributions.
  - Detecting trends in the comments and contributions, as well as spam and campaigns.
  - Calculating the sentiment of a particular comment or of all the comments and contributions.
  - Calculating similarity and relations between comments and contributions.
  - Producing statistics on the comments and contributions.
• Multi-lingual support.
• Possibility to reuse existing solutions and further build on them.
• In house expertise exists.
	• Investments are required for development and training.
	• It is not clear how existing solutions can scale to support more requests and increased demand/use.

Opportunities	**Threats**
• Synergies with DORIS of DG CNECT or other solutions used by Commission services. • Lead by example and develop added value services based on the exploitation of data analytics that help on the context of policy formulation. • Included in the Better Regulation toolbox.	• Technology limitations in terms of multilingual support. • There has to be caution not to reinvent the wheel. • The received contributions are not representative and hence cannot be used for drawing statically significant conclusions.

Investing further on text mining, as an enabler of stakeholder consultation analysis, will support policy officers with processing and analysing comments and contributions received in public consultations. This alternative will equip policy officers with tools that will allow them to have access to valuable information, hidden previously in the text, which they can utilise in the context of the policy lifecycle, thus improving the quality of the decisions taken and leading to time efficiencies. Overall, this alternative will have a positive impact on the accuracy of the analysis performed.

In order however to ensure the successful implementation of the alternative, the text mining techniques, algorithms and open-source solutions to be reused must be carefully assessed before being selected.

In order to achieve the desired outcome with a newly developed tool, we can divide the capabilities of the tool in several blocks that can be considered apart as their functionalities doesn't overlap or doesn't depend on the other blocs. A first block would handle the translation of the text if needed. The next step is to do some text processing to extract as much information as possible from the different comments or contribution, this step will include named entity extraction, sentiment analysis, natural language processing ... All the information extracted from the comment are now available to analyse, but a last step in order to ease the understanding of the data will be to perform some statistical analysis on the result as well as some clustering.

The information would be gathered from a survey tool in text format, several format are possible : CSV, TXT, XLS ... A first step is to translate the non-English comments, to spot those comments tools like libTextCat[3] or guess-language[4] can be used. In order to translate those comment to English, Google translate[5] or CEF Automated Translation[6] Building block are tools that can perform the job.

For the text processing part (entity extraction, sentiment analysis...), the open-source software GATE[7], Apache OpeNLP[8] or the Stanford NLP[9] libraries (already used by DORIS) can be used or the already existing DORIS tool can be scaled up and used more widely. KNIME[10] is another open-source platform to consider for the technology selection.

The last step for the analysis of the result, the clustering can be done by the DORIS tool itself (which use the Lingo3G commercial algorithm), or libraries for comparing similarities provided by the Apache Foundation, or the Carrot2[11] open source framework (using Lingo algorithm, less effective but free) can be used for clustering.

4   SOLUTION DESCRIPTION

4.1   Benefits

The main benefits of the proposed solution for stakeholder consultation analysis tools are indicated below.

For public administrations:

- Automation of the analysis process via the:

    - Ability to identify the different opinions expressed in the comments and contributions;

    - Ability to detect trends in the comments and contributions, as well as spam and campaigns.

    - Ability to elicit the sentiment of a particular comment or of all the comments and contributions;

    - Ability to discover relationships between comments and contributions;

    - Ability to easily produce statistics on the comments and contributions;

    - Ability to process multilingual content.

---

[3] http://software.wise-guys.nl/libtextcat/

[4] https://bitbucket.org/spirit/guess_language

[5] https://translate.google.com/

[6] https://joinup.ec.europa.eu/community/cef/og_page/catalogue-building-blocks#AT

[7] https://gate.ac.uk

[8] https://opennlp.apache.org/

[9] http://nlp.stanford.edu/software/

[10] https://www.knime.org/

[11] http://project.carrot2.org/

- Improvement in the flow of information from the citizens into the policy lifecycle;

- Improvement of citizen engagement in policy making;

- Realisation of time and resource efficiencies in processing and analysing comments and contributions received in public consultations.

For stakeholders (consultation contributors or any individual or organisation with an interest in the topic):

- Ability to participate in a transparent consultation process for policy making which avoids an error-prone and possibly biased analysis.

4.2 Success Criteria

The high-level success criteria of the proposed solution are

- Policy officers can quickly have a summary report of a consultation, outlining the key opinions and arguments, their sentiment and their strength with a level of accuracy of 80%.

- The amount of consultations published by the Commission, which receive above 1000 responses and whose results are analysed using text mining is increased by 50%.

- There is a reduction of 75% in the time that policy officers spend in processing, analysing and visualising the comments and contributions received in public consultations.

- There is a reduction of 75% in the human resources that are busy with processing, analysing and visualising the comments and contributions received in public consultations.

4.3 Assumptions and constraints

It is assumed that the solution will process only EU official languages.
It is assumed that the implementation of the solution will start with a well scoped pilot, before doing any large scale software development.
It is assumed that there exists enough interest within the EU institutions and the Member State administrations to participate in such a pilot.

4.4 Scope

The scope of the solution is limited to the development of open-source text mining solutions for analysing stakeholder consultations, focusing on the aspects outlined in the "Expected Outcomes".

The solution will be used at corporate level within the Commission.

The expected outcomes will be covered by one tool from a set of tools DIGIT intends to develop to support Participatory knowledge for supporting decision making. This set of tools can be included in the Better Regulation Toolbox, launched by the European Commission's Secretariat-General. For additional information, we refer the reader to *Appendix 1: References and Related Documents.*
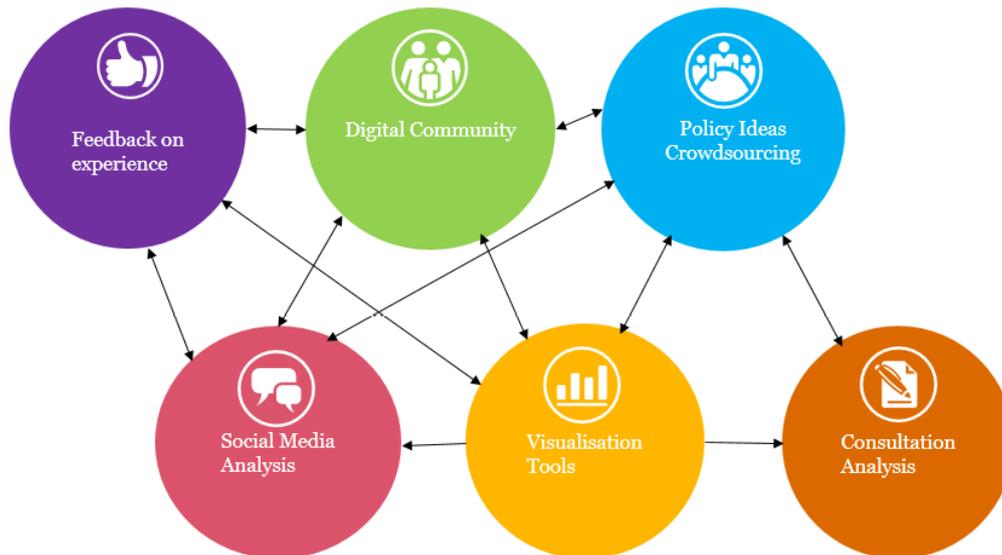
Business cases

Figure 1: Links between the business cases for corporate solutions supporting participatory knowledge for supporting decision making.

4.5 Solution Impact

The adoption of Alternative C: Develop text mining solutions for analysing stakeholder consultations is expected to have a significant positive impact on the "Conduct consultation work" phase of the proposed open public consultation process under the Better Regulation Guidelines, as detailed below:

- **Analysing stakeholders' collected input and feedback.** The existing stakeholders' consultation analysis process will be automated and consistent across the European Commission DGs and/or within a public administration in a Member State. The tool will be used on all stakeholders' consultations run by the European Commission or by public administration and the feedback received will be integrated into the policy-making process in a cost-efficient, fully reliable and transparent way.
- **Establishing a time- and resource-optimised process.** The process for analysing open questions will be optimised, as the tool will display the key elements of the contributions received. The increased number of stakeholders' consultations' contributions will therefore not significantly impact the time needed for their analysis and then subsequent integration into the policy-making cycle.
- **Need for a stakeholder consultation analysis team.** The proposed solution introduces the need for establishing a stakeholder consultation analysis team who is responsible for managing the solution and providing as a service to different Commission DGs or other organisations. The solution would therefore drive the development of a corporate service within the Commission DGs.

The solution will positively impact the stakeholders involved (including residents, citizens and organisations), as follows:

- **Trust.** The stakeholder consultation analysis tool will provide transparency into how the information collected will be analysed before used in policy making.

Summarising, the impact of Alternative C: Develop text mining solutions for analysing stakeholder consultations on both organisation and process is high.

4.6 Deliverables

The main deliverables of the proposed solution is a software solution. However, the adoption of a new solution will imply the development of additional deliverables such as:

- User manuals;
- Technical documentation;
- Branding strategy, including a communication plan; and
- Training material.

4.7 Risks

The main risks derived from the implementation of the proposed solution are the following:

- Owners of existing solutions will not agree to use the new solution, despite the limitations faced by their own solutions. This risk could be mitigated by conducting informative sessions highlighting the benefits (efficiency, time-saving, errors' decrease) driven by the usage of the tool.
- The criteria for the analysis of the comments and contributions are not well-defined, hence the outcomes of the analysis are of dubious quality. This risk could be mitigated by conducting training sessions on how to use the analytics tool.
- There is organisational resistance from people who feel that are losing control or ownership, because the processing and the analysis of comments and contributions is now automated. This risk could be mitigated by conducting informative sessions highlighting the difference between the analysis (automated) and the evaluation phases of the collected feedback in the policy-making cycle.

If the mitigation measures are not effective, the risks can have an impact on a low perception of the long-term benefits and overcome of these risks might require additional time and means.

4.8 Costs, Effort and Funding Source

The costs, effort and funding source will be defined by the business case owner. The business case owner of the stakeholders' consultation analysis is DIGIT.

The internal tasks currently foreseen are the following:

- Service management

The current estimation of internal resources needed for the management of the service is 1 FTE per year.

The external tasks currently foreseen are the following:

- Project management
- Business analysis & requirements
- Development
- Testing
- Evolutive maintenance
- Training
- Change management activities
- Operation

The current estimation of external resources needed for the development and operations of the service is 2 FTEs per year.

It is foreseen that the resources needed for the stakeholders' consultation analysis operations will be reassessed on a yearly basis, in alignment with the actual demand.

4.9 Roadmap

The current roadmap for the development of this business case is the following:

- The service description to be completed in 2016;

- the tool development and implementation to take place in 2016 and 2017; and

- the service operations to start between 2017 and 2018.

4.10 Synergies and Interdependencies

In the European Commission, some synergies can be found among the initiatives that are being carried out. Three main initiatives are identified:

- DORIS tool, currently owned by DG CNECT for their (and other specific DGs) stakeholders' consultations analysis, including semantic analysis and data visualisation capabilities.

- LEOS - Open Source software for editing legislation, a pilot that ran until Q4-2015 under ISA 1.13 action which tool enables the management of several versions of the same document (e.g. each version updated by a different stakeholder) for collecting stakeholders' feedback on draft legislation.

- KOEL tool, currently used by DG FISMA and DG GROW for their EUsurvey stakeholders' consultations analysis.

- The Better Regulation toolbox, launched by the European's Commission Secretariat General as a complement to the Better Regulation Guideline presented in in SWD(2015) 111.

5 GOVERNANCE

5.1 Project coordination

The DG that will coordinate the project is DG Informatics (DIGIT); together with the ISA2 Committee. DIGIT can nominate external suppliers to develop the technical part. However, DIGIT will be accountable for the delivery of the final solution and the support for its well-functioning.

5.2 Service Provider

The DG that will make the final solution available to other DGs and/or Member states is DG Informatics (DIGIT).

Within the Commission, the solution will be provided as a corporate service run by DG Informatics.

5.3 Associated services/stakeholders

The associated services will take part in the definition of the requirements, the performance, the guidance and/or providing lessons learnt of the project.

The identified associated services for this project are the following:

- DG CNECT;

- DIGIT; and

- SG.

A list of additional potential associated services and stakeholders for the project implementation have been identified in the above section *Synergies and Interdependencies*.

5.4 Beneficiaries

The beneficiaries of the project will be the European Commission DGs. During the project, the beneficiaries will be represented by the ISA2 Committee.

APPENDIX 1: REFERENCES AND RELATED DOCUMENTS

ID	Reference or Related Document	Source or Link/Location
1	D02.01 Participatory knowledge for supporting decision making Study Final Report	https://webgate.ec.europa.eu/CITnet/confluence/x/zgFWH
2	Better Regulation toolbox	http://ec.europa.eu/smart-regulation/guidelines/docs/br_toolbox_en.pdf
3	D03.01: Business cases for corporate e-Participation solutions	https://webgate.ec.europa.eu/CITnet/confluence/x/0AFWH