Project acronym:	BYTE
Project title:	Big data roadmap and cross-disciplinarY community for addressing socieTal Externalities
Grant number:	285593
Programme:	Seventh Framework Programme for ICT
Objective:	ICT-2013.4.2 Scalable data analytics
Contract type:	Co-ordination and Support Action
Start date of project:	01 March 2014
Duration:	36 months
Website:	www.byte-project.eu

Deliverable D5.1:
The BYTE Vision

Author(s): George Papachristos, *Delft University of Technology*
Scott W. Cunningham, *Delft University of Technology*
Claudia Werker, *Delft University of Technology*

Dissemination level:	Private
Deliverable type:	FINAL
Version:	1.4
Submission date:	29 February 2016

Table of Contents

Executive Summary 4

Three trends shaping the big data policy agenda 4

External Forces which shape Big Data Decision-Making 5

Scenarios that Support Planning 5

Sector Specific Visions for the High-Technology, Open Data Future 6

Common Forces Leading to a Common Policy 6

Comprehensive and complete BYTE Vision 6

1 Introduction 8

1.1 Abstract 8

1.2 Purpose 8

1.3 Organization 9

1.4 Policy Analysis Templates 10

1.5 Big Data as an Arena for Policy Action 10

2 Research Design 12

2.1 Requirements 12

2.2 Evaluation of Methodological Options 14

2.3 Method Outline 16

2.4 Workshop 18

3 Cross-Impact Analysis 20

3.1 Review of Cross-Impact Methodology 20

3.2 BYTE Requirements for Cross-Impact Analysis 21

3.3 Novel Extensions to Cross-Impact Analyses 21

3.4 Procedures for Gathering Cross-Impact Data 22

3.5 Cross-Impact Analysis Results 25

3.6 Trends Emerging from Cross-Impact Analysis 28

3.7 Review of Content Analysis and Monitoring methods 29

3.8 Justification for Emerging Trends Given Content and Monitoring 30

3.8.1 Big data transition 31

3.8.2 Big data hegemony 31

3.8.3 Big data regime 32

4 Futures Wheels 33

4.1 General Use of Futures Wheel 33

4.2 Specific Use of Futures Wheels for Interaction 34

4.3 Specific Use of Futures Wheels as Input for Cross-Impact Analysis 36

4.4 Specific Use of Futures Wheels for Brainstorming 37

5 Scenario Analysis 40
5.1 History and Purpose of Scenario Analyses 40
5.2 Overview of Scenario Approaches 41
5.3 Specific Foresight Methodologies 44
5.3.1 Manoa School Scenario Approach 44
5.3.2 Field Anomaly Relaxation 46
5.4 Scenarios in the BYTE Vision 48
5.5 Sector Specific Visions 52
6 Relevance for Other Sectors 53
6.1 Potential Full Relevance for Big Data 53
6.2 Grouping of sectors by externalities 55
6.3 Judgmental approach for scoring new sectors 55
7 Discussion and Conclusions 57

Tables

Table 1 List of method input, intermediate and final output 14
Table 2 A list of strengths and weaknesses of the three scenario methods 15
Table 3 List of methods and related requirements. 16
Table 4. Workshop agenda 19
Table 5. Magnitude of cross-impact vectors 25
Table 6. Cross-impact results 26
Table 7. Workshop agenda for Futures Wheels 35
Table 8. Categories of scenario techniques (adapted from Borjeson et al., 2006) 43
Table 9 Standard Industrial Codes (SIC) 54
Table 10 Classification of cases under SIC 55

Figures

Figure 1. Major points raised by the BYTE vision 4
Figure 2. Four BYTE scenarios 6
Figure 3. Objectives of the workshop 18
Figure 4. Definitions of impact 20
Figure 5. Novel requirements from BYTE 21
Figure 6. List of external forces 23
Figure 7. Blank cross-impact table 23
Figure 8. Criteria for rating impacts 24
Figure 9. Cross-impact table 24
Figure 10. Futures wheel examples (adapted from Curry and Schultz 2009) 34
Figure 11. Futures wheel template 36
Figure 12. Additional external forces 38
Figure 13. BYTE project futures wheel 39
Figure 14. Manoa school scenarios and logic 49
Figure 15. BYTE scenarios and cases 50
Figure 16. Continued growth vignette 51
Figure 17. Steady state vignette 51
Figure 18. Transformation vignette 51
Figure 19. Collapse vignette 52
Figure 20. Further relevant sectors 56

EXECUTIVE SUMMARY

The following document contains the vision statement for the BYTE project. The BYTE project is a multi-disciplinary and multi-sectoral study investigating the social, economic, political and legislative impacts of big data in Europe. The goal of this vision document is to set forth a comprehensive policy and capability planning framework for Europe, based on big data externalities revealed through the BYTE project case work. The five major points for discussion in the BYTE vision are given in the boxed text below. Appropriate sections for additional reference are given in parentheses. Each of these points is discussed in the appropriate subsection of the executive summary which follows.

1. There are three trends shaping the big data policy agenda for Europe: the transition, hegemony and regime of big data (3.8).
2. An extensive set of external forces which shape the context of big data decision-making are described (4)
3. The BYTE project has created a set of graphical and narrative scenarios to better support planning (5)
4. Interviews suggest that certain European sectors are particularly poorly equipped to address a set of high-technology, open data futures (5.5)
5. Many sectors in the big data economy face similar external forces, and can benefit from a set of common policies for data (6.2)
6. Although the BYTE case work is comprehensive, there are specific sectors where additional attention is necessary (6.3)

Figure 1. Major points raised by the BYTE vision

THREE TRENDS SHAPING THE BIG DATA POLICY AGENDA

You may read more about the document and its purpose and organization of this document in section 1, the

Introduction. The BYTE project vision argues that three major trends are shaping the big data world. Policy-makers and big data stakeholders should be informed of these trends, and be making necessary plans to hedge and shape European futures in the public and the private sphere. These trends are the big data *transition,* the big data *hegemony,* and the big data *regime.*

The big data transition describes the speed and extent to which big data technologies are adopted by European industry and government. Europe should be prepared for a rapid transition to big data, as well as the possibility that the transition is failed, forestalled, or otherwise delayed. The big data hegemony describes the extent to which big data technologies are controlled by a few big governmental or industrial actors. Decision-makers should be prepared to negotiate for a positive future for Europe in a world where a few key actors set the tone for big data. Likewise, European decision-makers should also be ready to participate if big data is governed by many small players, as could occur for instance in a vibrant marketplace of data, vendors and software. The big data regime describes whether big data will be governed in a system where privacy and proprietary knowledge is protected. Alternatively the regime may move towards a more open sector where data is treated like a public good. Both data regimes may have merit for Europe if decision-makers are prepared for the change. These trends and their justification are discussed more fully in section 3.8 Justification for Emerging Trends Given Content and Monitoring.

EXTERNAL FORCES WHICH SHAPE BIG DATA DECISION-MAKING

These three trends of transition, hegemony and regime are well-founded within an extensive body of BYTE case work. This case work is integrated in a single framework suitable for the development and analysis of policy. You may read more about these policy analysis templates in section 1.4, the Policy Analysis Templates. The templates reveal thirteen external forces which our BYTE respondent reveal are occurring again and again across sectors. Eighteen respondents are asked to exercise their judgment about the co-incidence and severity of these external forces.

The relationships between these forces are analysed in a rigorous manner by means of a well-established foresight technique known as cross-impact analysis. This technique integrates diverse perspectives on the future, and also weights and sorts the priorities of various stakeholders. You may read more about the cross-impact analysis technique and the novel extensions required for the BYTE project in section 3.4 Novel Extensions to Cross-Impact Analyses and section 3.5. Procedures for Gathering Cross-Impact Data. Or you may also read more about the statistical procedures which underpin the findings that the big data transition, hegemony and regime structure the external environment for the BYTE case studies. This is documented in section 3.6, Trends Emerging from Cross-Impact Analysis.

Consistent with the goals of the BYTE vision, this document also asks whether the case study work was comprehensive enough in identifying possible new externalities affecting the European public and private sector. This question is answered through an extensive brainstorming exercise with BYTE stakeholders identifying an additional six external forces affecting the big data future. You may read more about these forces in section 4.2, Specific Use of Futures Wheels for Interaction. You may also read more about the futures wheel technique which was used as a basis for this systematic examination of both forces and externalities. This is documented in section 4.1 General Use of Futures Wheel. Futures wheels are also used as a basis for finding higher order and emergent forces in the big data future. This is described in section 4.4, Specific Use of Futures Wheels for Brainstorming. A new and emergent set of externalities is revealed by a close look at the transition, hegemony and regime trends of big data.

Scenarios that Support Planning

The three big data trends are embedded in a set of scenarios. Scenario analysis techniques prepare decision-makers for the future by anticipating a broad and varied set of possible future circumstances. The objective of scenario analysis techniques is not to predict the future, but to instead prepare decision-makers by anticipating a range of possible circumstances. Policy-makers and decision-makers can thereby make choices which are well-adapted to a range of possible futures. Policies can be developed to shape the future, maximizing the benefits received from beneficial externalities. Policies can also be developed to hedge against futures, offsetting any negative externalities which can come to pass. The vision statement presents a set of four scenarios, presented both in narrative and graphical format. These scenarios follow a lesser-known technique known as the Manoa school of scenario analysis. This scenario technique is particularly well-suited for analysing the big data future given the open and porous character of big data technologies.

You may read more about the BYTE scenarios in section 5, Scenario Analysis. You may also wish to read more about the methodological foundations of this and other scenario analytic methods in section 2, the Research Design. Regardless, a quick visual representation of these scenarios is provided in the figure below.

Sector Specific Visions for the High-Technology, Open Data Future

Despite the fact that a vision of European big data futures is now growing clear, not all sectors are equally prepared. Europe may be particularly prepared for a set of high-technology and open futures. Such futures are described by some as post-capitalist, and these futures emphasize the creation of economic value in a world where data is now omnipresent. For these futures the crisis management and environment sectors may be particularly poorly prepared given the revealed externalities of the case. You may read more about these issues in section 5.5 Sector Specific Visions.

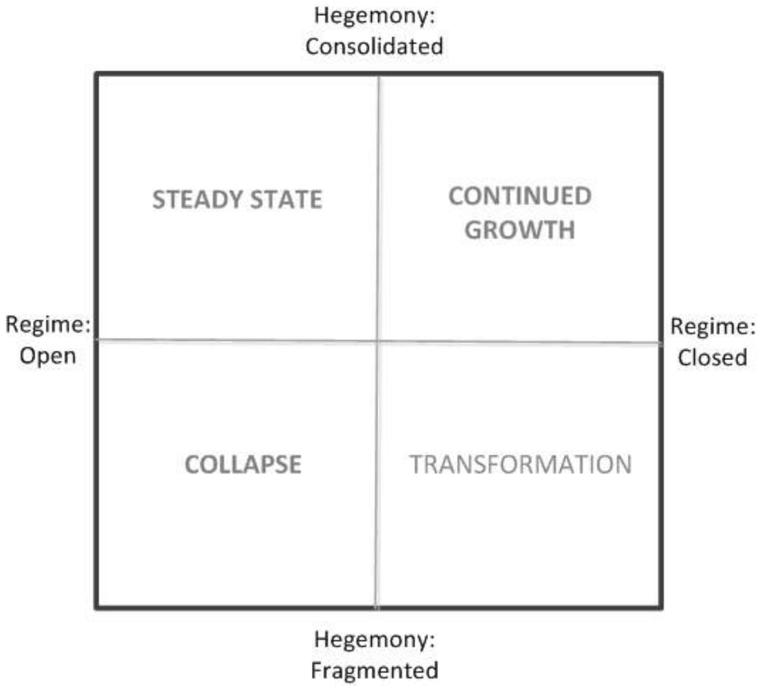


Figure 2. Four BYTE scenarios

COMMON FORCES LEADING TO A COMMON POLICY

The BYTE project now has a comprehensive vision on both the external forces shaping big data adoption, and the externalities created by the adoption of big data. The creativity and brainstorming exercise demonstrate that a sufficiently rich context which describes external forces has been developed by means of previous BYTE case work. This document argues further that cross-case learning is possible, particularly for sectors concerned similar external forces. Such sectors may benefit from a common structure of legislation and governance. You may read more about this argument in section 6.2 Grouping of sectors by externalities.

COMPREHENSIVE AND COMPLETE BYTE VISION

Another goal of the vision statement is to evaluate the comprehensiveness of the BYTE case studies, and to consider how and if BYTE policy recommendations can extend beyond the cases considered by the project. The BYTE cases are remarkably general across many industrial sectors, as classified by the standard industrial code system. Nonetheless there are two sectors in particular, which are already anticipating a post-capitalist future, which need particular attention to be incorporated into future BYTE project recommendations and dissemination. These are the infrastructure, insurance, and finance sectors. You may read more about these conclusions in section 6.3 Grouping of sectors by externalities. This foreseen gap will be further addressed in the roadmapping work package (WP 6), as well as in the big data community (WP 7) and dissemination packages (WP 9).

1 INTRODUCTION

1.1 ABSTRACT

This document and deliverable presents the vision of the BYTE project. In the eyes of our analysts and advisory board two overweaning trends are shaping the future of big data. The first trend identified by our panel is the *hegemony* of big data. Who will control the future of big data – the many, or the few? The second is *regime* of big data. Will big data ultimately be governed in an open, and cooperative fashion, or will big data become a closed and proprietary regime? At the intersection of these trends we create four possible scenarios for the future. This deliverable uses these scenarios to enhance planning, and to extend the vision to sectors previously unconsidered in the BYTE project.

1.2 PURPOSE

The BYTE proposal outlines four related tasks for work package 5. Task 5.1 involves producing sector-specific visions for big data. A series of 6 sector specific visions will be built for the year 2020 based on the cases that were developed and detailed in WP3. Project partners will identify big data challenges that are visible at present and influence the potential positive and negative externalities that can be captured in particular sectors. This will provide the input for a gap analysis used in the road mapping in work package 6 that will examine what knowledge, technologies, applications, policies and capacities are necessary in order to support the effective development of big data.

Task 5.2 is to produce a single, general vision for big data, where the positive externalities are amplified and the negative ones controlled. It will utilize deliverable 4.1 (horizontal analysis of cases) as input. Part of task 5.2 consists in identifying big data challenges around capturing positive externalities and diminishing negative externalities across different contexts, more broadly than in task 5.1. Thus, 5.1 will be partially used as input to 5.2 but its broader scope implies that additional work will be needed to "fill in the gaps" of a general big data vision for 2020 that will include positive and negative impacts (externalities).

Task 5.3 involves specifically considering what new externalities may be brought into being as a result of the achievement of the big data visions produced in Tasks 5.1 and 5.2. It requires considering 1st and 2nd order impacts of the impacts that are currently visible, but also new that may emerge in the near future. Task 5.3 is anticipatory is anticipatory in character and to an extent it will utilize work done for 5.1-2 in two ways. Exploring future cross impacts of existing positive and negative impacts, and future impacts of trends and drivers of big data documented in 5.1-2.

This document and deliverable has five major purposes:

1. Present the BYTE vision
2. Describe how stakeholders and advisory board members are incorporated in the vision
3. Generalize the findings across sectors
4. Anticipate new impacts
5. Generalize the findings to new big data sectors

The first and primary purpose of this document is to present the BYTE vision. The vision is a Europe which is adaptive and agile in the face of an uncertain future. In pursuit of this vision the document presents four very different scenarios of the big data future. The scenarios build upon and extend previous case work in the BYTE project. In particular the scenarios incorporate the external forces, or key uncertainties, previously identified in the BYTE case studies.

The four constructed scenarios enable European actors, stakeholders and decision-makers to become better prepared regardless of which future actually transpires. The scenarios are created by BYTE project members, by expert participants, and by the BYTE advisory board. A secondary purpose of this document is therefore to describe how these disparate views of the future are synthesized into a consistent and prioritized view of the most pressing trends affecting European citizens, government, and industry.

A third purpose of the document is to generalize the findings of previous case study work. The scenarios and underlying trends of the BYTE vision are a natural vehicle to compare and contrast across the seven case studies of the project. The purpose is therefore to share learning across seemingly disparate sectors, and to prepare and develop recommendations suitable for all sectors regardless of their specific orientation to big data.

A fourth purpose of the document is to anticipate heretofore unanticipated externalities and impacts. The BYTE vision provides a higher level perspective on seemingly disparate forces affecting the European big data ecosystem. Given this higher level perspective it becomes much more feasible to integrate diverse externalities, and to anticipate emerging forces. Generalizing the BYTE project requires that we compare across the BYTE cases, grouping the sectors according to related impacts. This grouping of cases also serves as a mechanism by which we can potentially extend the results of the BYTE project to entirely new sectors not fully considered by our cases.

1.3 Organization

This deliverable consists of a document with six major sections. The sections are the research design (section 2), cross-impact analysis (section 3), a futures wheel approach (section 4), scenario analysis and sector specific visions (section 5), and a discussion of BYTE for additional potentially relevant sectors (section 6). These can be further grouped by an over-arching design (section 2), the treatment of the foresight analysis techniques (sections 3 – 5), and the treatment of sectoral issues (sections 6). The document concludes with a problem statement which makes claims for policy action in light of BYTE research activities to date. Each of these sections is described in more detail in the paragraphs which follow.

There are three foresight analysis techniques utilized in this deliverable. Cross-impact analysis is a technique which may be used to further analyze the externalities previously identified in the BYTE project. A futures wheel approach is a graphical technique for examining the interaction of externalities, and eliciting feedback from experts and laypeople alike. Scenario analysis techniques are an appropriate technique for synthesizing disparate forces and presenting a unified vision.

Each of these sections outlines the technique, and applies the technique to the BYTE project. A brief methodological summary is provided to help assure reproducibility of results. Readers uninterested in methodology may safely omit these sections.

As noted there are two sectoral issues to be addressed in the deliverable. The first entails integrating and unifying the disparate externalities identified in the BYTE project. A unified perspective on the externalities better enables the individual sectors and cases to prepare for an uncertain future. A study is presented here whereby seven unforeseen futures is presented.

A second goal is an exploratory analysis which examines the relevance of the BYTE project to new sectors, not yet investigated by the project. The externalities and impacts previously identified serve as a basis for this sectoral extension to the BYTE findings.

1.4 Policy Analysis Templates

The material to follow is dependent on a set of policy analysis templates. These templates generalize and extend the BYTE case study materials as described in deliverable 3.2. These templates are discussed more fully in deliverable 5.2, but are briefly introduced in this section for the first time.

The case study work presented in the BYTE project describes seven cases of big data usage, drawn from seven European sectors. One of the objectives of deliverables 5.1 and 5.2 is to derive general lessons from these cases, and to look ahead to general policy recommendations. For both these reasons deliverable 5.2 takes a perspective drawn from policy analysis. The discipline of policy analysis uses engineering principles to systematize and structure problems, as a step towards supporting effective decision-making. As a policy analysis framework of choice, we utilize the XLRM framework (Lempert et al 2003).

The XLRM framework organizes analysis efforts around a particular system or domain for study. A system is compromised of various subsystems and the various relationships between the subsystems. By knowing these relations (the R in the framework) we can help decision-makers design and select appropriate policies. The framework also considers the critical decisions, or policy levers, with which decision-makers can affect and steer the decision. This is the L in the XLRM framework. Monitoring the framework requires that we examine various concrete measures and outcomes of interest from the system – this is the M in the XLRM framework. Despite the best laid plans of decision-makers, various external forces or wildcards disrupt policy and planning. This is the X in the framework.

The BYTE project adds one additional element to the framework – the various actors in the system. Whereas traditional policy analysis is developed with a single actor in mind, it is increasingly necessary to create plans for a polycentric environment, where there are multiple actors with divergent needs and capabilities. This adds the fifth or final element to the policy framework. The resultant modified framework is called the A-XLRM framework, where A stands for the variety of actors which, in this case, inhabit the European big data ecosystem.

The A-XLRM framework is used to inventorize actors, external forces, policy levers, system relationships and measures. The framework is applied across all seven of the BYTE cases, as discussed in deliverable 5.2. The resultant input is an important precursor for the foresight activities in both deliverable 5.1 (this document), as well as deliverable 5.2.

1.5 Big Data as an Arena for Policy Action

Although this document endorses the use of policy analytic techniques to support decision-making, big data is still a challenging topic for policy analysis methods. There are three reasons. The first is that the system boundaries are ill-defined. The second is that there is a variety of different actors in the system, all with varied interests and capabilities. The third is that the analysis is evolving and foreward-looking.

The term big data is a phenomena or an epiphenomena, but it is not necessarily a single policy domain. A recent social history of the term suggests it was first used by computer hardware companies in the Bay Area of California. The term was originally used as a placeholder to signify these companies' general interests with market finding. Indeed these companies and others did find their market and as a result there is an increasingly widespread use of computers through out society. Even despite this pinpointing of the origins of the word, it is inherently difficult to draw a boundary around the topic of big data. The BYTE vision solves this problem through a process of induction, fed by ample empirical evidence from the various BYTE cases.

A second challenge is that there are multiple actors capable of taking decisive action in the European ecosystem. Given the mandate of the BYTE project, we must also consider key stakeholders as well – often the European citizen. In order to better address this challenge the BYTE vision uses techniques for framing the problem of big data, and for understanding the most critical trade-offs of big data. Much of this material will be addressed in deliverable 5.2.

The BYTE project is forward-looking, tasked with looking for major externalities for the next five years. Many policy arenas are also forward-looking. Nonetheless the field of big data is unusually complex from a technical as well as a social perspective. Technically, the field is characterized by combinatorial innovation. Although the various components of big data (hardware, software, data, domain and market) may be individually well-known, the combination of these factors creates a fundamental incompleteness about the rules of the game. Regardless of the technical uncertainties in the area, the social and legislative uncertainties of the field may be still greater. The BYTE vision employs classic planning techniques, including scenario analysis, to address these challenges.

As will be further discussed and justified in the sections below, the BYTE project anticipates policy problems needing further corrective action at the European Commission level and at the level of member states. Case study work to date acknowledges a range of sector specific impacts, but the work has not noted why these impacts are a problem requiring policy action. An actual problem statement requires a gap analysis which states how and why action in the big data space is not measuring up to European needs and requirements. Given workshop activities and cross-impact analysis (the work of which is further reviewed below), the BYTE project anticipates that corrective action is needed in three specific areas.

The first area involves the move towards the big data transition. European policies must be timed to do no harm if the transition is forestalled, yet be swiftly moving enough to take advantage of new opportunities. The second area involves the hegemony of big data. Europe must grapple with the potential presence of big players in the space of big data. Alternatively, Europe must be equipped to deal with the fragmentation and disorganization which could result from a diverse ecology of big data players. The third arena involves the regime of big data. Europe must be equipped for a forthcoming era of open data and sharing, while still creating the necessary incentives to forestall underinvestment in common infrastructure and resources. Alternatively, Europe must anticipate and act decisively to create a regime of strong data ownership.

This document is not to predict a single future for Europe. Nor is it to forestall legislative decision-making. Given these goals the document will set forth a range of possible visions for the big data future. These visions can be used to develop robust policies for Europe regardless of what specific futures come to pass. Furthermore these scenarios can be used to develop policy scorecards which can be used by European decision-makers to select the appropriate measures.

The next section discusses the first of the major foresight techniques used in the BYTE project – a cross-impact analysis. The section describes the foundations of the technique in the literature. Several extensions of the technique, custom developed for the BYTE project are then outlined. The section highlights the major external forces of big data, as identified by the various BYTE case studies, and generalized through the use of policy analysis templates. These serve as a basis for individual and group exercises to explore the future of big data. The resultant trends are then described.

2 RESEARCH DESIGN

Given the novel requirements set for the BYTE project, an innovative research design will be needed for the forsight activities of work package 5. This section sets forth a multi-methodoology which combines various methods to meet the challenges of the project. The section consists of the requirements section (2.1) which recapitulates the requirements of the project. Section 2.2 enumerates specific foresight methods, and evaluates these methods in light of the BYTE project requirements. Section 2.3 sets forth a multi-methodology which combines the best aspects of various existing techniques. It then outlines the necessary steps to conduct the multi-methodology. The concluding subsection describes a stakeholder workshop which is an integral component of the workpackage.

2.1 REQUIREMENTS

The primary input for task 5.1 is the detailed case study report delivered in task 3.2. This deliverable outlines the stakeholders identified and interviewed for each case, the data sources, uses, flows, the main technical challenges, and provides a big data assessment. It also discusses economic, social, ethical, legal and political externalities present in each case. This concludes the documentation of each case.

The input for task 5.2 is the horizontal study delivered in task 4.1, which provides a big data assessment and an overview of the technical challenges involved in each case study, and then discusses common, economic, social, ethical, legal, and political impacts. It discusses big data practices in the case studies along the range of technical challenges through which big data is often characterized: volume, velocity, variety, and veracity. It then compares the societal externalities in the cases and checks which externalities are common and which are unique to each sector.

The value of task 5.2 is that it utilizes the work done in task 4.2 that cuts across the big data case studies in 3.2. The deliverable of 4.2 makes obvious that there are *no* common positive or negative impacts (except innovation) across the cases, and some cases do not have anyimpact types at all. Therefore task 5.2 will need to go beyond that, "fill in the gaps", for example by considering additional core sectors that were not documented in 3.2 or 4.2 in order to build a coherent societal picture of how big data will influence society in the future.

A more careful consideration of the methods and a final choice requires an explicit statement of requirements drawn from work package 5 description. Tasks 5.1-3 are considered in stating method requirements, as is the need to maintain continuity with prior work carried out in the BYTE project and outputs produced up to this point and the subsequent tasks in the remaining work packages.

Owing to the unique nature of the BYTE project, the work package description and concomitant requirements, it is possible that a single method will not satisfy all the requirements. This implies that the possibility of combining some steps from different methods should also be examined. This possibility is explored after the evaluation of a range of foresight methods.

1. Incorporate diverse casework: This is a requirement derived from by having the tasks 3.2 and 4.1, as inputs for tasks 5.1 an 5.2 but also from the nature of the issue under consideration. Big data is an all-encompassing term and it is likely that it will have an impact across different sectors. This requirement applies to both tasks 5.1 and 5.2 owing to the diversity of cases examined in 3.2 which constitutes the input for 5.1, and the scope of task 5.2 which requires consideration of a diverse range of sectors altogether.
2. Incorporate best practices and policy levers. This requirement comes from the assessment of the main externalities associated with the use of big data in task 4.2 and the identification of ways – best practices to reduce or mitigate the negative effects and amplify the positive externalities. Task 4.2 identifies a range of 6 best practices that can have a cross sectoral but also sector specific impact. They may not apply in their entirety to the cases examined in 3.2, thus the requirement is not as strong for 5.1 as it is for 5.2 where it is anticipated that all the best practices identified in 4.2 are relevant. Thus this requirement is relevant for tasks 5.1 and 5.2.
3. Maximize beneficial impacts; minimize negative impacts. This requirement is specified by the original BYTE proposal.
4. Provide suitable output for work package 6. This is a requirement derived by the work breakdown structure and scope of the BYTE project. Ideally, the output of work package 5 will be part of the input for work package 6. Under this requirement the intermediate steps of each method and the output produced there will also be considered as potential input for work package 6 which involves developing road maps. This is a particular requirement for task 5.2.
5. Handle temporality of visions. A concomitant requirement relating to work package 6 is how each method handles the temporal aspect of visions i.e. whether the method lend itself to producing output that can be used as an input for a roadmapping exercise.
6. Constructively engage with stakeholders in a workshop. The core of the method should lend itself to application in one or two day workshop where participants can have some meaningful input. This implies that it should be possible to introduce the method to the participants and conduct relatively short sessions where they apply parts of the method and provide content for work packages 5.1 and 5.2. This requirement applies to tasks 5.1 and 5.2.
7. Incorporate diverse world views. The case studies in task 3.2 demonstrated the wide ranging impacts that big data can have. This implies a wide range of stakeholders and a range of different perspectives through which these impacts can be seen as positive or negative. Hence, it is necessary to consider the impacts documented in tasks 3.2, 4.1, and 4.2 through a range of different viewpoints that accommodate the world views of likely stakeholders. It follows that this requirement applies both to task 5.1 which is sector specific and 5.2 which is cross sectoral in its scope.
8. Identify emerging big data impacts. Tasks 5.1-2 involve building visions for big data both sector specific and cross sectoral based on documented big data impacts in tasks 3.2 and 4.2. However, it is likely that new big data impacts will arise in the future as a result of moving towards and achieving the big data visions produced in tasks 5.1 and 5.2. The method then applied in work package 5 should facilitate the exploration of unanticipated impacts emerging in the future both in the output of tasks 5.1 and 5.2.

2.2 Evaluation of Methodological Options

In this section we provide a reasoned evaluation of why we choose a specific methodology or methodologies. The choice of method for carrying out tasks 5.1-3 requires looking into existing available methods and their characteristics. It is also important to consider the input and output of each method for carrying out tasks 5.1-3. This is because the method that will be applied in work package 5 must utilize output from prior work packages produced in the BYTE project as input, and produce relevant output that can be utilized in subsequent work packages and tasks as input, in particular work package 6. The choice of the method applied in tasks 5.1-3 will be based on the consideration of inputs required for each method, their intermediate steps and outputs, the strengths and weaknesses, and the extent to which each method covers requirements relevant to tasks 5.1-3.

A starting point is an overview of the methods considered with an overview of inputs, intermediate steps and outputs (Table 1). Methods considered are the Manoa, Field Anomaly Relaxation (FAR) and Conflict Analysis (and more?). Then the strengths and weaknesses of the methods are summarized (Table 2) before a statement of what are the method requirements for tasks 5.1-3. Finally, an assessment of the methods is done and the steps that will be followed are laid out in detail.

Table 1 List of method input, intermediate and final output

Methods	Input to method	Intermediate output corresponding to method steps	Final Output
Manoa School	Critical trends	1. Construction of futures wheel. 2. Long-term consequences of present trends. 3. Cascading consequences of present trends	Four scenarios: Continued growth, Collapse, Stable, Transformation.
FAR	Alternative futures for the domain of concern.	1. Aggregate description of the future into a number of factors. Levels define the range of possibilities in each factor. 2. When a number of factors are combined they produce a field e.g. with 2 factors of 2 levels each, a field has 2x2 possibilities. 3. Logically inconsistent fields are eliminated or "relaxed". 4. Those remaining are combined into distinct, different plausible time sequences to produce scenario outlines.	A set of comparably plausible scenarios for a chosen field, each descriptive of changing circumstances (rather than events) over a future span of one to three decades.
Conflict Analysis	Knowledge and preferences of actors.	1. Analysis of actor moves and their combinations. 2. Specification of payoffs 3. Recommendations based on strategic behavior	Conflict graphs representing specific outcomes and the moves available to actors.

The choice of a method for carrying out tasks 5.1-3 will be based on the requirements, the strengths and weaknesses of each method. This should provide some insight into the range of requirements that each method is likely to satisfy. Drawing on this, the possibility of combining elements and steps from different methods will be explored to ensure that all of the requirements are finally met.

Considering the Manoa approach, its main strength is that it has a heuristic allowing for considerable creativity. This lends itself nicely to adjusting the method to the particular tasks of the work package and also the planned workshop (task 5.4). In this way it is possible to pursue the analysis of particular impacts to a satisfactory level of detail. On the other hand it is not as rigorous as Field Anomaly Relaxation (FAR) and therefore it does not produce a similar depth of analysis overall.

FAR in contrast is more rigorous and this enables the integration of information into a set of alternative futures from which eventually a smaller set is produced. On the downside the number of scenarios can easily grow very large and the method does not produce the most probable or plausible scenarios.

Table 2 A list of strengths and weaknesses of the three scenario methods

Method	Strengths	Weaknesses
Manoa School	• It is a creative and effective method at producing vivid, provocative detail and generating different outcomes across all levels of a possible future. • It facilitates understanding of the dynamics of change rippling through various systems, as drivers create primary, secondary, and further cascades of impacts, which then create cross-impact turbulence.	• Unlikely to generate the depth of analysis of the Causal Layered Analysis or FAR approaches. • Output depicts the future of society.
Field Anomaly Relaxation	• It rings scattered information and insights together, allowing composition of a smaller set of alternatives from which internal inconsistencies have been removed. • It allows for a rigorous process that considers all possibilities and is traceable.	• The number of scenarios can easily grow. • It does not derive the most probable or plausible futures. • Potentially time consuming- Several iterations for anomaly relaxation might be necessary.
Conflict Analysis	• Finding and resolving social dilemmas. • It makes clear how actors arrive at their choices and if they reach a collective dilemma. • The method is very adaptable to desk research, strategic advice settings and workshops.	Difficulty in valuation of real world outcomes.

Table 3 indicates the requirements satisfied by each method taking into consideration the respective strengths and weaknesses of each method. The requirements are not disaggregated per task as it is judged that there is sufficient overlap. A mark ∨ is assigned to a method if it satisfies the particular requirement more than others do or if it is the only one satisfying it. The right most column indicates whether a single method satisfies all requirements and can thus be applied in its entirety in tasks 5.1-3.

This is obviously not the case. The picture that emerges from Table 3 is that a combination of some methods is required for these tasks. In particular a synergy between the Manoa School of scenarios and Field Anomaly Relaxation seems to be possible provided that the steps of each method are appropriately combined. This is because they appear to facilitate the delivery of complementary aspects of tasks 5.1-3 while not having a considerable overlap.

Table 3 List of methods and related requirements.

Method	1	2	3	4	5	6	7	8	Requirements met
Manoa	✓	✓				✓		✓	4/8
Field anomaly Relaxation		✓	✓	✓	✓				4/8
Conflict Analysis	✓	✓	✓	✓					4/8
Intersection	✓	✓	✓	✓	✓	✓		✓	8/8

A particular requirement that none of the methods meets to a satisfactory degree is requirement 7: the ability to incorporate diverse world views of the stakeholders involved. This is necessary as different stakeholders with opposing world views will likely take opposing positions over particular big data impacts. Their stance needs to be at least acknowledged and considered in laying out road maps in work package 6 because it will enable the identification of potential opposition and obstacles for the road maps. It follows that this is a particularly significant requirement that cannot be ignored in carrying out tasks 5.1-3.

2.3 Method Outline

In summary, the discussion has provided an overview of the methods considered for tasks 5.1-3, their strengths and weaknesses, and has revealed the need to combine elements of the Manoa school and FAR in order to come up with steps that will fulfill all the requirements stated earlier. In addition culture theory will be used at the last step in order to structure the thinking on the likely attitudes of diverse kinds of stakeholders. In laying out the steps work package 5 will follow we choose to start with the Manoa school as FAR begins with imagined futures as its input, the end outcome of the package.

1. The first step consists in listing influential trends, technical, societal and business challenges and construct futures wheels to trace first and second order cascade impacts of big data. These may or may not be overcome within the time horizon set for the vision. Thus, they may or may not enable the amplification of positive externalities and the control of negative ones. A careful consideration of the wheels should identify overlap of impacts coming from more than one wheel. These will need to be tabulated. 1st and 2nd order impacts from each wheel are tabulated and this should produce possible future combinations of impacts. It is likely that they will be numerous. The implications of

overcoming the technical challenges in terms of the impacts identified for the stakeholders in each case should be worked out.
2. In the second step, step 3 from FAR will be used to summarize and produce a set of uniquely distinct, plausible set of impacts that constitute parts of future visions.
3. The third step borrows from the Manoa and involves incasting the sets of future visions into the four archetypal scenarios that Manoa proposes. An assessment of the impacts will produce at the very least a list of gains and losses for the stakeholders. In line with methodology detailed in task 3.1 impacts are considered with respect to: (i) business models, (ii) data sources and open data, (iii) technologies and infrastructures, (iv) policies and legal issues, and (v) social and ethical issues. Then this will be the input to the scenarios generated following the Manoa school. These will be further elaborated and any inconsistencies will be put up for discussion with consortium stakeholders in a workshop.
4. The fourth step will involve considering the range of views stakeholders are likely to hold for each big data scenario. This involves considering and classifying the range of stakeholders involved in work packages 3 and 4 under the four categories of the culture theory in order to produce aggregate positive and negative stakeholder evaluations of the scenarios produced.

There is an explicit requirement for developing visions in work package 5. We interpret this to require the development of scenarios where the positive impacts are amplified and the negative ones reduced. This requires that the widest possible group of stakeholders has something to gain rather than something to lose. It is expected that different scenario types will approach more to a best-case scenario for each of the six cases documented in task 3.2. For example, the best-case scenario for the smart city case can be of the transformation type, while for the energy or environment cases it can be of the growth type. We don't anticipate that the collapse scenario will constitute the basis for the best case scenario in any of the six cases. The range of scenarios produced will be systematically cross pollinated with the culture theory stakeholder categories.

For WP5.2 a single, integrative vision is required that is cross cutting all case studies and ideally would be informed by the work produced in D5.1. Here, an effort will be made to draw the lines between the two deliverables. Nevertheless, there are two constraints: (i) all impacts do not manifest across all case studies therefore it is not possible to consider them in the aggregate, and (ii) the context of each case is generally disjunct from another. While an effort will be made to use the work in 5.1 the vision what will be developed in 5.2 will be based on an aggregate assessment of the challenges, and impacts detailed in task 4.1. If necessary we will go back to the case studies detailed in task 3.2.

It is harder to come up with some definitive stakeholder list fort task 5.2, unless we define ad hoc the broader public sphere and EU citizens as the core stakeholders. This is not necessarily inconsistent with the definition of stakeholders of task 5.1. In fact, if while producing the vision becomes possible to cast each case specific vision in terms of the broader public then this will go a long way towards delivering task 5.2.

An additional input for the aggregate vision in 5.2 is deliverable 2.1 the report on big data issues. This provides an overview of economic, legal, social, ethical, and political issues that are related to big data that are non-specific to the cases reported in D3.2. A second input is appendix B in deliverable 3.1 which lists positive and negative impacts in terms of the interactions between three kinds of stakeholders: public sector, private sector and citizens. These impacts will be the basis for developing the single vision. As with deliverable 5.1 we cannot anticipate which type of scenario will come closer to the best case and thus to a big data vision, although it is unlikely that the collapse scenario will be a candidate. Then, the steps to construct the single vision follows the same steps of deliverable 5.1.

In task 5.3 we will consider what new externalities may be brought into being as a result of the achievement of the big data visions produced for 5.1 and 5.2. A direct link exists with these deliverables and it is the application of the Manoa methodology. A central part of this is the futures wheel (Figure 1). This will be used to map and develop thinking about the long-term consequences of existing trends (A,B,C) and the potential emergence of new externalities as a result of them continuing in the future. These trends will have been already identified in 5.1 and 5.2.

2.4 WORKSHOP

A significant component of work package 5 is a stakeholder workshop. This subsection outlines the design and conduct of this workshop. This material will be re-visited and reincorporated in the subsequent sections in this document of cross-impact analysis (section 3), foresight wheels (section 4), and scenario analysis (section 5).

The workshop was a full day exercise, with twenty-five attendees. The attendees were a mix of academic, non-governmental and industry participants. The main objectives of the worksho, as provided to attendees, are given in the figure below.

You should attend the workshop if you'd like to:

1. review and systematize the impacts of Big Data across six sectors
2. examine the critical assumptions underlying European Big Data governance
3. anticipate the need for the privacy enhancing design of systems and services
4. participate in creating four visions of Big Data, with milestones for tracking and action
5. discuss opportunities for action in hedging and shaping the Big Data future

Figure 3. Objectives of the workshop

The first two objectives are met through policy analysis techniques. These are more fully described in deliverable 5.2. These constitute sessions 1 and 2 in the workshop (see table below).

Table 4. Workshop agenda

Time			Topic	Mode
8:30	9:00	Registration and Introductions		
9:00	9:15	Agenda	Briefing	
9:15	10:45	Session 1. Review of Previous Findings	Moderated Discussion	A systematic review of the Big Data issues identified in seven sectors.
10:45	11:00	Coffee Break		
11:00	12:00	Session 2. Problem Structuring, Root	Moderated Discussion	Reviewing the critical assumptions and underpinnings of European governance.
12:00	13:00	Lunch		
13:00	13:50	Session 3. Vision/ Futures	Group Work	Specifying four future visions of European Big Data, with milestones for action.
13:50	14:40	Session 4. Scenario Creation	Group Work	Discussing actions to hedge and shape our European Big Data future.
14:40	15:00	Coffee Break		
15:00	15:45	Session 5. Keynote Soeech	Speech	Anticipating the need for privacy by design in Big Data systems and services.
15:45	16:00	Next Steps and Conclusions	Briefing	

Objectives 3 and 4 are part of the foresight activities as described in this document. Both sessions center around a futures wheel technique (described further in section 4). The fifth and final objective of the workshop was a keynote speech on privacy by design.

This session was added because it was topical, and because it is beneficial to have a variety of different activities in a workshop design. The speaker was Prof. Jeroen van den Hoeven, a professor of the philosophy of technology at the Delft Unviersity of Technology. The professor spoke about the needs for privacy-enhancing, and socially responsible design in big data technologies. The speech presented a number of thesis statements resulting in an active discussion by participants.

In this section a multi-methodological foresight activity is designed to meet the requirements of the BYTE project. Two components are needed – an *ex ante* policy analysis, and a foresight methodology. The ex ante policy analyais activity, conducted in the morning of the workshop, is more fully documented in deliverable D5.2. Consideration of diverse stakeholder needs and requirements, and an explicit policy problem statement, are particular hallmarks of such *ex ante* analyses.

The foresight activity involves a modified scenario analysis technique originally developed in Manoa, Hawaii. The complex and porous boundaries of big data activities demand that the project use a more narrative scenario approach. The resultant approach combines cross-impact analyses, scenario planning, and stakeholder workshops. The role of these methods is discussed more fully in the sections below.

3 CROSS-IMPACT ANALYSIS

The following subsection is a brief review of a prominent foresight and forecasting technique known as cross-impact analysis. The purpose of the technique is to gain an improved understanding of the nature and structure of uncertainties affecting a decision. Uncertainties may be symptoms of deeper, underlying causes.

3.1 Review of Cross-Impact Methodology

A few definitions are needed to clarify terminology. The BYTE project discusses externalities. The economic concept of externalities is originally intended to capture the concept of goods or bads not directly captured by the market place. The BYTE project often uses this terminology more generally to indicate the unintended consequences of a big data technology. A long literature on technology impact assessment uses the term impact to captur the idea of unintended social, economic, political or legislative consequences.

Impacts are subject to the choices made by decision-makers to amplify or alleviate consequences. We must therefore clearly distinguish between external forces – or events outside the capacity of decision-makers to easily affect – and technological impact. One may Of course the distinction between impact and external forces may not always be entirely clear since the decisions made by one actor in a problem domain may induce consequences for another. The technique of cross-impact analysis, discussed further below, often confounds external forces and impacts. As will be further discussed, impacts often compound and interact. We use the term trend to signify a constellation of impacts, some of which have already been observed or experienced, and others which may still be emerging.

- External Force: This is an external event, unforeseen by decision-makers.
- Externality: Used in an economic sense, this indicates additional value or cost of a good or service which has not been valued by the marketplace.
- Impact: This is the unintended consequence, used for instance when assessing a new technology. The term may also be used to signify the costs of a decision which is borne by others.
- Trend: A constellation of impacts which occur together, as revealed by estimates of conditional probability.

Figure 4. Definitions of impact

Cross-impact analysis originally emerged out of early business simulation games. Many of these early pen and paper games used physical randomizers like cards and dice as a way of confronting business decision-makers with real-world uncertainties. Such uncertainties served as a vehicle for players encounter some of the complexities of real-world decisions. Effective play entails making decisions which are robust to a variety of unforeseen circumstances.

The original creators of these games moved their paper games online, simulating the random events using computers. Given the conditional structure of events and probabilities, cross-impact tables were created. The general purpose application of these tables even outside of simulation gaming was soon realized. A more generically applicable cross-impact analysis procedure was thereby developed.

Current cross-impact methodology entails creating a finite list of possible external forces and uncertainties. This list becomes the row and the column of the resultant cross-impact matrix. The resultant cells are interpretable as conditional probabilities. For instance, at the intersection of row A and column B the resultant cell is interpretable as the conditional probability of event A occurring given the occurrence of column B.

A coarser representation of probabilities can be used, for instance by scaling the probabilities using a five point scale. The elicitation of probabilities can be performed with naive users or experts alike, particularly when the simplified representation of cross-impact probabilities are used. The resultant mathematical analysis uses Markov chain analyses, and closely related techniques eigenvalue analyses, to identify the limiting probabilities of the distribution.

3.2 BYTE Requirements for Cross-Impact Analysis

The underlying structure of impacts, as identified by cross-impact analysis techniques, reveals one or more underlying trends. It is helpful to identify these trends since, while decision-makers may be well aware of some impacts, they may be ill-prepared to recognize and act upon other follow-on events. Thus cross-impact analysis assists in decision-making since it makes the full panoply of emerging trends much clearer for analyst and decision-maker alike.

Cross-impact analysis is an appealing approach for use in the BYTE project. A variety of external forces are identified through case study work. The policy analysis templates (described in the introduction, and further detailed in deliverable 5.2) afford an overarching view across all cases and all the identified external forces. Cross-impact analysis can help unify and integrate these impacts, perhaps helping to identify whether these impacts have common cause. The resultant trends can then be incorporated in plans to hedge risk, or shape future opportunity.

However the BYTE project sets forth a set of additional requirements for cross-impact analyses. The technique must be usable with a variety of different stakeholderrs, project members and experts. The technique itself involves a fairy tedious procedure of analyzing many cells in a cross-impact matrix. The task of eliciting probabilities needs to be modified to permit discussion and interaction within a workshop setting. The technique must integrate potentially disparate visions of the future. Furthermore the technique must prioritize the variety of underlying forces so that the most significant forces can be identified, and incorporated into planning. The BYTE vision, even for relatively short-term planning, requires that we consider external events as uncertainties, not as risks.

3.3 Novel Extensions to Cross-Impact Analyses

Seven modifications to a standard cross-impact procedure are adopted in response to the requirements of the BYTE project. These are briefly outlined, and more fully described in the section below. These modifications are summarized in table 2, below, and then discussed more completely below.

1	Incorporation of positional information
2	Incorporate external forces rather than impacts
3	Lack of prior probabilities
4	Symmetric matrices representing co-occurrence of events
5	Five point reduced factor scales
6	Use of eigenvalue analysis to structure trends
7	Use of eigenvalues to prioritize the most important factors

Figure 5. Novel requirements from BYTE

A wider variety of perspectives on the future were sampled during the foresight workshop associated with the BYTE vision. For this workshop it was important to incorporate a more interactive and graphical procedure to elicit probabilities. These probabilities were incorporated into the main cross impact analysis results. More detail about these procedures are described in section 3, futures wheels, discussed below.

The cross-impact procedure is run using external forces previously identified from the BYTE case studies. This is an easy modification; cross-impact studies use a variety of different uncertainties as input. Furthermore these external events are taken as uncertainties, rather than risks. Therefore the likelihood of their occurrence is not measured using probabilities, or incorporated as prior probabilities in the analysis. The resultant outputs are therefore more suited for dealing with a deeply uncertain future. The results are incorporated into the scenario analysis, which is discussed more fully in section 4, below.

The BYTE cross-impact study uses the probabilities of co-occurrence, rather than the conditional probabilities of events. This simplifies the greatly simplifies the elicitation procedure, and reduces the burden of analysis and interpretation. The results are congruent with, if more limited than, the full cross-impact analysis.

Further simplifications involve the use of a five point Likert scale for analyzing the mutual occurrence of events. These are then analyzed using eigen analysis techniuqes. As a result of the analysis a set of eigenvectors are calculated. The resultant vectors correspond to the underlying trends in the system, plus a superfactor. The superfactor serves as a set point for the scaling of the probabilities.

Each eigenvector has a corresponding eigenvalue, which corresponds to the importance or weighting of the trend by the cross-impact participants. The resultant weights are interpretable as a distributed voting procedure, where the most heavily weighted trends show the greatest concern and the greatest coherence in assessment across the raters.

In the next subsection the procedures for collecting the cross-impact data is discussed. This involves discussing the external forces selected for analysis, the interview protocol used with experts, the input of futures wheel information, and the calculation of underlying trends and their assigned priorities by experts.

3.4 PROCEDURES FOR GATHERING CROSS-IMPACT DATA

The cross-impact analysis is based on the external factors as identified in the policy analysis templates. The templates are discussed more fully in deliverable 5.2. A unified set of external forces are presented below, for convenience in the discussion.

Europe is subjected to . . .

A. extreme climactic or environmental change, including events such as the melting of the arctic pass.
B. a series of unusually severe environmental problems and inclement weather.
C. increasing resource scarcity resulting in geopolitical conflict
D. demographic changes resulting in a less-trained, lower capacity workforce outside of Europe.
E. greater diversity in culture and language
F. dominant players in information technology infrastructure. This may involve national actors, such as the United States, or a concentration of market power among key industrial providers of technologies and services.
G. increased reluctance to share data, perhaps derived from intellectual property law and the extension of copyright.
H. increasing amounts of environmental pollution generated by actors outside Europe.
I. demographic change in the rest of Europe resulting in older and sicker populations.
J. catastrophic outbreak of disease in the rest of the world, perhaps spreading to Europe.
K. enhanced innovation in public health and epidemiology, generated by research institutions all across the world.
L. enhanced innovation in urban data and infrastructure management, generated by research institutions all across the world.
M. a severe rash of cybercrime and the criminal misuse of data.

Figure 6. List of external forces

As shown in table 3, thirteen external forces were found across all cases. Note that by definition all external forces are outside the control of European actors, although European actors may choose to act to enhance or counter the forces. These external forces, as identified by the BYTE case studies, involve a mix of social, environmental, political, legislative and technological forces.

The forces are arranged in a cross-impact matrix, and cross-impact participants are asked to rate the relationships in the matrix (figure 1). The cross-impact matrix is implemented in Excel. Participants either rate the impacts directly in Excel, or are given a printout of the spreadsheet and template.

		A	B	C	D	E	F	G	H	I	J	K	L	M
A	Climate Change, Environmental Change, Artic Pass, Inclement Weather	5												
B	Natural Disaster, Environmental Problems		5											
C	Geopolitical Resource Conflict			5										
D	Demographic Change: Lower Human Capacity				5									
E	Linguistic Diversity					5								
F	U.S. Dominance in Infrastructure, Big Players in IT						5							
G	External closed policies on data ownership and copyright							5						
H	Environmental Pollution								5					
I	Demographic Change: Sicker, Older Populations									5				
J	Disease Outbreak										5			
K	Innovation: Epidemiology											5		
L	Innovation: Urban Data												5	
M	Criminal Misuse of Data													5

Figure 7. Blank cross-impact table

When using the table, participants are asked to rate the co-occurrence of the external forces using a five point scale. See the text box below for the specific instructions. The matrix is symmetric, so the upper triangle need not be completed. In addition the diagonal need not be rated. Filling out the complete cross-impact matrix requires entry of 48 scales.

> Please fill in the following cross force matrix with your own personal assessment of the occurrence of the two rated forces. You may want to consult the force listing when thinking about these co-occurrences. Some forces have an opposite.
>
> 5. When one force occurs, the other occurs highly
> 4. When one force occurs, the other is slightly more likely to occur
> 3. When one force occurs, the other may or may not occur
> 2. When one force occurs, the other is slightly less likely to occur
> 1. When one force occurs, the other is much less likely to occur
>
> In short, are the forces mutually reinforcing, or ultimately generated by the same underlying causes?

Figure 8. Criteria for rating impacts

Ultimately three analysts rated the cross-impact matrix. An additional fifteen ratings were obtained from workshop participants using a futures wheel approach. The futures wheel approach is described in more detail in section 3, below.

The complete ratings of participants are archived in project repositories. Our goal here is to provide a quick, visual inspection of the raw data, and to provide more interpretive findings regarding the underlying trends. This is shown in figure 2, below. In this heat chart representation of the raw data, green values represent forces which are highly likely to occur together. Red values are forces which are less likely to occur together, or are countervailing. So for instance, our participants thought forces A and B (climate change and natural disaster) are more likely to be coincident.

		A	B	C	D	E	F	G	H	I	J	K	L	M
A	Climate Ch	92,5	87,9	59,2	58,0	56,2	30,6	34,0	75,0	58,6	62,0	61,3	55,0	37,1
B	Natural Di	87,9	92,8	57,4	59,1	57,9	29,8	34,0	75,9	58,7	63,8	65,8	59,1	43,5
C	Geopolitic	59,2	57,4	82,3	57,1	36,0	52,9	52,9	45,4	58,7	47,9	41,3	50,3	54,3
D	Demograp	58,0	59,1	57,1	62,7	52,7	58,3	51,2	59,3	59,3	59,7	58,6	53,7	58,5
E	Linguistic I	56,2	57,9	36,0	52,7	74,5	50,8	37,4	59,3	48,4	53,7	60,3	60,9	52,6
F	U.S. Domi	30,6	29,8	52,9	58,3	50,8	91,1	73,0	47,3	51,9	42,3	43,8	54,2	65,1
G	External cl	34,0	34,0	52,9	51,2	37,4	73,0	73,6	45,0	47,8	44,8	39,6	40,8	49,2
H	Environme	75,0	75,9	45,4	59,3	59,3	47,3	45,0	82,6	58,9	62,8	64,9	59,2	47,1
I	Demograp	58,6	58,7	58,7	59,3	48,4	51,9	47,8	58,9	61,9	62,3	58,2	59,3	55,1
J	Disease Ou	62,0	63,8	47,9	59,7	53,7	42,3	44,8	62,8	62,3	83,5	78,6	70,0	56,5
K	Innovatior	61,3	65,8	41,3	58,6	60,3	43,8	39,6	64,9	58,2	78,6	81,4	69,0	53,6
L	Innovatior	55,0	59,1	50,3	53,7	60,9	54,2	40,8	59,2	59,3	70,0	69,0	73,0	68,0
M	Criminal M	37,1	43,5	54,3	58,5	52,6	65,1	49,2	47,1	55,1	56,5	53,6	68,0	81,1

Figure 9. Cross-impact table

The cross-impact table is further analyzed inside Excel, using Excel Solver. The eigenvector problem involves the following optimization problem: Find a set of weighted orthogonal vectors which, when multiplied together, reproduce with minimum error the original data as rated by the participants. This is a standard expression of the eigen-analysis problem. It is somewhat unusual to use a non-linear optimization procedure to find the least squares solution to the problem, but the problem runs quickly and completely within Excel using Solver.

For this analysis four vectors are extracted. One vector is lost without generality, since the five point scaling could be replaced with any number of alternative ordinal rankings. Two trends are useful for scenario analysis. An extra trend is extracted for exploratory purposes.

3.5 Cross-Impact Analysis Results

The results of the cross-impact analysis is analysis is discussed further in this subsection. The main result of the cross-impact analyses are three trends which emerge from the concerns and priorities of experts and stakeholders. Summary results from the cross-impact analysis is given in the table below.

Table 5. Magnitude of cross-impact vectors

Code	Name	Eigenvalue	Variance	Cumulative Variance
R	Residual	1309,8	57,5%	57,5%
I	Intercept	741,4	32,6%	90,1%
T1	Transition	121,2	5,3%	95,4%
T2	Hegemony	40,2	1,8%	97,2%
T3	Regime	64,2	2,8%	100,0%

The table lists the three trends resulting from the analysis, an intercept, and a model residual. These model components are given by name and code. Statistical outputs including eigenvalue, and variance and cumulative variance are also given. The eigenvalues and variance show the relative weighting and priority of the trends as revealed by user ratings. These trends are further discussed and interpreted below.

The user ratings are both complex and noisy. The resultant cross-impact analysis explains some 42% of the variance in the user ratings, leaving an additional 58% unexplained by the model. A greater proportion of variance is explainable if additional factors are extracted from the model.A balance is sought in the model between robustness and reproducibility, and simplified results which can be graphically and narratively interpreted. As a result only the first three potential factors are extracted. This is appropriate given the purpose of the model, which is exploratory rather than predictive.

The intercept of the model is, as discussed, an arbitrary offset which is determined by the rating and scaling of the model. Alternative rating procedures can be introduced with no loss of generality. However little specific significance can be attached to the intercept rating. The remaining three trends are named transition (the most important), hegemony (the least important of the three), and regime. More detailed results exploring the content and resultant interpretation of these trends is provided in the table below.

Table 6. Cross-impact results

Code	Description	Absolute Scaling					Relative Scaling			
		Offset	Transition	Hegemony	Regime		Offset	Transition	Hegemony	Regime
A	Climate Change. Environmental Change. Artic Pass. Inclement Weather	7,92	-4,48	0,67	2,61		0,291	-0,407	0,106	0,412
B	Natural Disaster, Environmental Problems	8,12	-4,46	0,29	1,77		0,298	-0,405	0,046	0,279
C	Geopolitical Resource Conflict	7,05	1,48	-2,93	4,42		0,259	0,134	-0,462	0,698
D	Demographic Change: Lower Human Capacity	7,62	0,85	0,03	0,69		0,280	0,077	0,005	0,108
E	Linguistic Diversity	7,17	-0,48	2,57	-2,30		0,263	-0,044	0,405	-0,363
F	U.S. Dominance in Infrastructure, Big Players in IT	6,90	6,09	2,16	0,32		0,254	0,553	0,341	0,050
G	External closed policies on data ownership and copyright	6,24	4,52	1,59	1,94		0,229	0,411	0,251	0,307
H	Environmental Pollution	8,04	-2,22	2,59	0,33		0,295	-0,202	0,409	0,051
I	Demographic Change: Sicker Populations	7,54	0,32	-1,05	0,62		0,277	0,029	-0,165	0,098
J	Disease Outbreak	8,10	-1,24	-1,80	-2,36		0,298	-0,113	-0,283	-0,372
K	Innovation: Epidemiology	7,99	-1,63	-0,41	-3,04		0,294	-0,148	-0,065	-0,480
L	Innovation: Urban Data	7,90	0,25	-1,33	-2,53		0,290	0,023	-0,210	-0,400
M	Criminal Misuse of Data	7,31	3,37	-2,07	-1,80		0,269	0,306	-0,327	-0,284

The cross-impact table presents two sets of four columns. The columns correspond to the intercept (I), and three trends (T1,T2,T3) as discussed above. The two sets provide the cross-impact results in absolute number (left), and also in a unit scaling (right). The cross-impact model scales each trend by the magnitude of user ratings and concerns. It is therefore helpful to examine these trends both on an absolute as well as a relative scaling.

On the rows of the table are the thirteen external forces derived from the policy analysis templates, and previous BYTE case work. These forces are previously discussed in section 3.4. These external are given by code as well as by short interpretation. Interpreting the table requires an examination of individual forces, and their relative assignment to specific trends. Each of the trends is now interpreted in turn.

The offset column is, as previously noted, is uninteresting and is an artefact of user rating and questionnaire scaling. Most forces are given a comparable weighting with only minor variations seen. The next trend in the column is the transition trend.

This trend highlights factors such as U.S. dominance of big data infrastructure, and the presence of big multinationals in the big data space (loading 0.553). It is also associated with closed and proprietary policies for data (loading 0.411). Negatively scoring forces on this trend are climate change (-0.407), and natural disaster (-0.405). One interpretation of this factor is that it represents the general transition to big data, with the positive loading forces accelerating the trend, and the negative loading factors focusing money, public attention and policy away from the nascent transition. The other loadings are consistent with this interpretation.

The second trend, labelled hegemony, also involves a strong negative loading of geopolitical resource conflict (-0.462), and criminal use of data (-0.327). In the positive direction is environmental pollution (0.409), linguistic and cultural diversity (0.405), and the presence of big players (0.341). This trend is interpretable as an underlying force for structuring and ordering the system. The current hegemony permits a liberal trading regime, with varied cultural elements, that is supportive of many environmentally polluting industries. Users apparently anticipate the rise of conflict, disorder, and crime should this regime begin to disintegrate.

The third trend, labelled regime, shows a strong force for geopolitical resource conflict (0.698). Also heavily loading are climate change (0.412) and strong intellectual property and proprietary restrictions on the use of data (0.307). Negatively loading on this factor is innovation in the public health, epidemiology and urban planning realms (loadings -.480 and -.400). Also negatively related is linguistic and cultural diversity (loading -0.363). The trend can be characterized as open policies for data. The need for crisis management and climate adaptation are strong forces for coordination and planning. On the other hand innovative actors may want to keep a closed ecosystem, so that they can appropriate the benefits of their own research and development activity. Likewise local cultural identities and linguistic communities may want to close themselves off from unwelcome forces which impose standardization and uniformity.

This three trend model is a credible interpretation of the patterns of responses as revealed by users. Nonetheless the underlying, external forces may be fairly poor instruments for the conceptual and mental models of users. This suggests the need for a broader conceptualization of external forces. We take a deeper look at this in the subsequent section on the futures wheel.

An informal robustness test is applied to the result. The eigen structure of the trends is stable, and is not dramatically altered by adding or subtracting user ratings. The relative emphasis on trends does vary however according to the addition or subtraction of users. This suggests there may be relatively different priorities by users in assessing the future. Adding new users may change the comparative ranking in the three major trends, although this is unlikely to change the underlying structure.

3.6 Trends Emerging from Cross-Impact Analysis

The BYTE project foresees three major trends which are impacting all seven of the case studies. These trends are the technological transition towards big data, the hegemony and control of big data, and the governance regime for big data. The question underlying the technological revolution of big data involves the pace and extent of technological change, and the degree to which public and private attention and concern will focus on the virtual world instead of the natural world. The question underlying the hegemony of big data concerns the number and locus of actors, external to Europe, who will be involved in the governance of the technology. The question of the regime for big data concerns whether a regime of openness and sharing will emerged, or whether big data applications will be closed and proprietary. Ultimately this a question which concerns the appropriation and commodification of knowledge. Now that each of these trends has been outlined briefly, each of these trends are described in more detail in the paragraphs to follow.

One of the underlying trends in the BYTE vision is the big data revolution. Participants are concerned with the rapidity and extent of the big data transformation, regardless of whether this transition is for good or for bad. A number of external forces underlie this trend – including epidemics, chronic illness, natural disasters, severe environmental degradation, and exceptional problems of urban management. These forces serve as a focal point for future action. If big data provides actionable policies for addressing these challenges, big data is more likely to be part of a sustained transition with corresponding investment in public interest and policy development. Big data promises to disrupt existing ways of doing business. Thus, vested interests may wish to slow or hinder the adoption of the technology. The capacity of big data to provide actionable results for societal problems, and to disrupt existing ways of doing business, are themselves dependent on the underlying pace of technological change in big data technologies. Given this trend the BYTE vision must be adaptive to a variety of different transition paths towards the adoption of big data.

Another of the underlying trends in the BYTE vision is the hegemony and control of big data. European policy has a critical impact on the effective governance of big data, yet for planning purposes, Europe must also focus on those forces which are not in direct European control. The first force which must be addressed is a potential concentration of power. Conversely the future big data may be controlled and shaped by a multitude of actors. A second force concerns the locus of decision-making for big data. Critical decisions may be shaped in the public sphere, or these decisions may be controlled in the private sphere. These first two forces involving the concentration and locus of decision-making for big data are very general in scope. We can and should get more specific about the specific, critical actors which are influencing big data. Given this trend the BYTE vision must support planning for a variety of different governance arrangements for big data.

The third underlying trend in the BYTE vision involves the regime for managing big data. Big data may be managed as a public good, and may provide corresponding civic benefit. There may be efficiency and coordination gains by creating open repositories of public knowledge. But the governance challenges of managing open, standardized data sources may be made more severe by the scale of public big data initiatives. On the other hand, a closed regime presents its own set of costs and benefits. A set of closed and proprietary systems may provide a boost to innovation, as individual entrepreneurs are able to appropriate the benefits of benefits. A closed regime for big data imposes additional costs of potentially disparate and fragmented data sources. Private holders of data may claim a rent on their data, resulting in higher costs for the user. In truth an intermediate regime, neither open nor closed, is likely to result. Different regimes may be imposed in different sectors. Given this trend for big data regimes the BYTE vision must support a variety of different regimes for the sharing and the commodification of big data. In the next subsections we anchor these trends in actual societal discussion concerning the future of big data.

3.7 Review of Content Analysis and Monitoring methods

As noted in the previous subsection, three major trends were identified as a result of the cross-impact analysis. In addition to the hegemony and regime of big data governance, participants are also concerned with the rate of transition to the big data future. These trends will be further anchored in the literature and press on big data in section 2.7. Two qualitative analysis techniques are used to scan the literature – content analysis and monitoring. The purpose of this section is to provide a brief overview of these techniques.

Monitoring is the oldest and most general purpose of foresight analysis techniques. The procedure involves routine scanning of a variety of different information sources, including news articles, scientific articles, opinion pieces, and trade journal discussions. The purpose of this monitoring is to surface new and emerging trends and issues. The technique itself probably originated with the rise of modern bureaucracies in the 19th century. Researchers have argued that the technique was originally developed to meet a range of functional needs for these bureaucracies, including environmental management, colonial administration, and domestic counter-intelligence.

In its simplest and most basic form the technique involves the broad scanning of diverse sources of news and literature. The goal is to piece together a partial yet congruent picture of emerging issues of interest to government. In its original form the technique involves creating a dossier on topics of interest. Some monitoring analysts have a two-step approach whereby anything of general interest is first collected. The analyst then develops, by induction, a set of more coherent trends which is further used to direct and structure the monitoring activity. This monitoring approach is organized around paper print outs and manila folders.

This traditional intelligence technique was added to a suite of more modern approaches, first by RAND corporation in the 1950s. Thus, monitoring approaches were added to scenario analysis, simulation and modelling, expert opinion and Delphi, and the full suite of approaches now used in technological forecasting. This very old technique of monitoring is increasingly being updated to the 21st century by means of computers and big data. Broad-based scanning activities are conducted by computer, using bots. The actionable intelligence is then filtered by a mixture of machine learning and human judgment. Despite advances in automation, human judgment is still required even when computer monitoring is deployed.

The object of study in technology monitoring are usually trends. These are conceived of as broad, organizing principles which can be used to order and make sense out of seeming disparate environmental events. Given the emphasis on technological change it often makes sense for analysts to scan scientific articles, patents, popular science accounts, new product offerings, trade show announcements, and industry press. Social media is increasingly added to this list of sources as a way of monitoring societal participation and feedback.

Few changes if any are needed to adapt monitoring for big data purposes. Big data is an unusually broad, unusually diffuse object of study. Given this it is very approach to use trends as a basis for structuring monitoring activities. It becomes much more difficult to focus the monitoring activity since there are multiple sectors and technologies where big data activities are potentially relevant. There is an unusually broad societal participation in the big data discussion, making social and popular media a key element in monitoring big data. Marketing white papers are an important source of information. Big data is particularly topical and in the news, particularly given recent revelations about government monitoring and surveillance.

Content analytic procedures are an appropriate technique for ensuring the quality and reproducibility of any monitoring effort. These approaches were originally developed for application to written text. These approaches entail the quantification of content based upon reproducible measures as applied to text. These reproducible measures may involve coding exact words and phrases. Alternatively human judgment can be incorporated using more generic rules or schema.

The key validation procedures for content analysis involve testing for the content validity, construct validity, and statistical conclusion validity. Content validity involves demonstrating that a given text actually contains relevant content. Construct validity involves defending the idea that the material actually supports the conceptual schemes imposed upon it by the analyst. Statistical conclusion validity centers around defending the idea that a given content coding scheme can be consistently applied across raters. If there is high inter-rater reliability then the coding schema is presumably a valid measure of the actual content on the page.

3.8 Justification for Emerging Trends Given Content and Monitoring

We now discuss the three trends of the BYTE vision in light of societal discussion about the future of big data. The prelude discussing content analysis and monitoring provides necessary background context for understanding how a more qualitative approach of monitoring can complement the quantifiable and expert opinion oriented approach of cross-impact analysis.

The following discussion is based upon an monitoring effort on big data, conducted over the course of two years. The effort involved two analysts, spending three person-months of effort. The monitoring exercise primarily involved exploring social media sources, and secondary references embedded in these sources. Additional sources involved investigating conferences, text books, and the home pages of various organizations in both the public and private sector. The effort was primarily a manual approach, although some computer automated procedures were used. An acknowledged threat to validity for the monitoring effort is the need for reproducible coding schema, and the testing of these schema across a variety of raters. This threat is partially alleviated by the cross-validation afforded by the cross-impact analysis.

The following is a highly abbreviated account of the vast amount of discussion available on big data in the public and private media. The account is organized to provide the broadest possible perspectives on the big data phenomena. The three major trends found in the cross-impact analysis serves as an organizing principle with which to structure this material. This account is also partially documented in work package 2, which surveys the economic impact of new technologies in big data.

3.8.1 Big data transition

We begin the discussion with the trend of the big data transition. The big data transition involves consideration of the rapidity of technological change underlying big data. It also entails considering if big data will fully emerge, or will be forestalled by other societal concerns. There are a variety of different opinions about the rate and significance of these change in big data. Two technological optimists, anticipating very rapid change, are Ray Kurzweil and Robin Hanson.

Kurzweil, who is currently Director of Engineering at Google, has previously led a distinguished career as a computer scientist, futurist and inventor. Kurzweil describes exponential growth in computer related technology which ultimately leads to a singularity. The singularity, postulated to occur in the next twenty-five years, involves the ultimate merger of human and machine. Robin Hanson, an associate professor of economics at George Mason University, also presents an ambitious vision of the future. Hanson anticipates a new industrial revolution which will ultimately result in a hundred-fold increase in economic productivity. According to Hanson, this revolution has already begun, and should be fully manifested within the next two decades. Hanson underpins his arguments with long-term trends in human economic growth and productivity.

There are also big data skeptics. Cathy O'Neil is a prominent blogger, and in former careers had positions in academia and in quantitative finance. O'Neil argues that the big data revolution has been over-hyped. The adoption of the technology will ultimately be hindered by the inability to automate and simplify complex statistical analyses. O'Neil's scepticism harkens back to previous waves of hype and disillusionment which accompanied fields such as artificial intelligence, knowledge discovery, and data mining.

3.8.2 Big data hegemony

The next trend for discussion is the big data hegemony. Aral Balkan is an activist and entrepreneur living in England. Balkan presents a vision of big data which is dominated and consolidated by large companies effectively operating in consort with national intelligence and security aparatus. For Balkan the emergence of big data is a threat to democracy. A similar account is provided by Jaron Lanier. Lanier is an American computer scientist and composer. Lanier argues that the technological structure of big data ultimately leads to an insurmountable advantage for the first movers in the market. The result is a crowding out of the creative class in favor of those operating a few, very large "siren servers."

There are relatively few proponents who argue that big data is truly pluralist – either a democratic institution, or part of a fully competitive marketplace. Some computer scientists have researched the democratic ideals of the early internet, and the extent to which these visions and structures have persisted over time (Chung et al 2016). On the democratising side, arguably one of the most influential agents is Gene Sharp, founder of the Albert Einstein Institution. Sharp develops strategies for advancing freedom through non-violent action. The handbooks of the institutions have been instrumental in multiple pro-democracy movements and rallies across the world. Sharp's work is not inherently big data in character, although it relies on social networks, viral communication, and multi-lingual translation. Another instrument for democratic change, which does rely on big data is the United Nations PULSE laboratory. This laboratory, with offices in New York City, Jakarta and Kampala, harnesses big data for humanitarian aid, and uses big data to evaluate the policy concerns of average citizens. The PULSE laboratory is both democratising as well as open; an open regime is discussed more fully below.

As noted earlier part of the trend of the hegemony of big data involves consideration of whether big data will diffuse to, and benefit, small to medium sized enterprises. Intuit, an American tax software company, develops data solutions and targets them to small and medium enterprises. Intuit calls this process the democratization of big data. There are undoubtedly many other companies in fast moving sectors which hope to market big data solutions to smaller customers.

3.8.3 Big data regime

A final consideration is the big data regime, and whether data will be a source of closed and proprietary knowledge, or whether it will be an open resource for the public good. Many commentators have argued that the business model of the internet is fundamentally based on advertising. A similar argument could be made for some sectors of big data as well. One expression of a closed and proprietary regime is the vision of big data, and the internet, as a walled garden. The most prominent advocate of a vision like this has been Facebook.

As an example of a walled garden Facebook's Mark Zuckerberg recently offered free internet service to rural India. However the access was limited through a a small suite of services. The result is a closed internet where all sites and online sources are no longer equally accessible. India's Telecommunications Regulatory Office ultimately rejected the offer.

Facebook has been a particular target of activist Max Shrems. Shrems initiated two lawsuits alleging that Facebook violates European privacy laws by transferring personal data to the United States and directly to the United States National Security Agency. Ireland is the European headquarters of Facebook, and therefore the source of potential citizen redress. Although the case was rejected by the Irish Data Protection Commissioner, the case was heard by the Irish High Court and ultimately referred to the Court of Justice of the European Union. In the most recent events surrounding the case, the Safe Harbour framework, a diplomatic foundation of cross-Atlantic data transfer, was determined to be inconsistent with European law and therefore declared invalid.

Closed, proprietary data may also be a critical source of innovation (Redman, 2013). Proprietary data sources afford organizations a sustainable and competitive advantage. Companies are thereby incentivized to invest in innovation since they are more likely to be able to appropriate the benefits of their efforts. These investments may result in spillovers resulting in benefits for society as a whole. Legal scholar Frank Pasquale has described the need for adequate transparency and governance of algorithms, which are increasingly becoming a walled off domain of government and corporate activity (Pasquale 2015).

Perhaps the most radical vision of open data is given by the science fiction author and physicist David Brin. Brin argues that the big data transition has already progressed to a point where it can no longer be revoked. The answer to the dangers of big data is radical transparency – everyone knowing everything about everyone. Brin calls this regime sousveillance, or monitoring from within. According to Brin this policy of openness also promotes democracy since it permits citizens to monitor the powerful. Another researcher and institution of note in the space of open data is Max Roser at the University of Oxford. He uses data and effective visualization to instrument progress towards development and inclusive growth across the world.

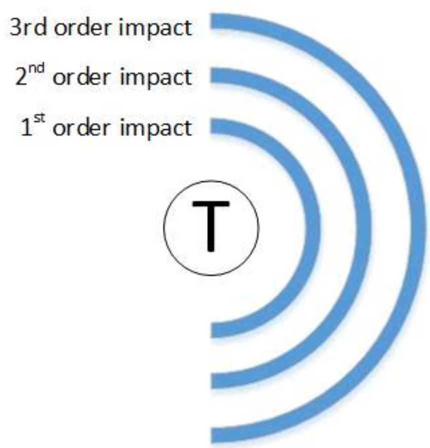
In the following (sections 3 and 4) we expand on these three trends by means of the futures wheel technique and scenario analysis methods. The futures wheel technique, discussed next, provides three actionable results for the BYTE project. First it enables a workshop-based activity for discussing potential big data futures. Second it allows these impacts to further feed into and enrich the cross-impact judgmental analysis. Third and finally the technique enables the BYTE project to anticipate new, and potentially unforeseen externalities.

4 FUTURES WHEELS

The futures wheel technique was invented as a means of thinking about the direct and indirect consequences of societal impacts (Futures Wheel 2009). Its inventor Jerome Glenn first developed the technique in 1971. The technique has also proven useful as an element in multi-methodological futures studies, as well as an effective vehicle for use in visual communication and synthesis.

4.1 GENERAL USE OF FUTURES WHEEL

Futures wheels seek to map and develop thinking about the long term consequences of today’s issues. The first way is to construct a futures wheel that maps out the implications of each of the trends considered. In a futures wheel, implications are mapped for each of the trends, placing the trends in the center. The trends and their interactions are considered as a form of wemerging issues analysis when constructing of futures wheel. The two figures below are examples of futures wheels which have been adapted from Curry and Shultz (2009).



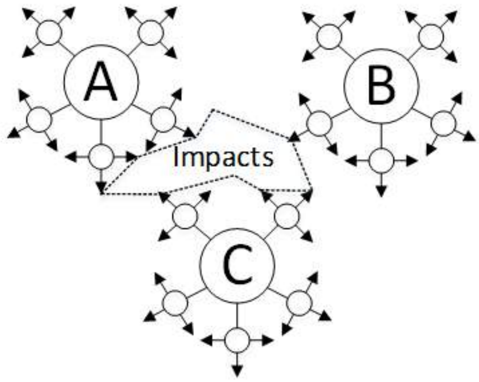


Figure 10. Futures wheel examples (adapted from Curry and Schultz 2009)

As an example, the influence of a particular new technology 20 years from now. The futures wheel does not quit after finding only the first order impacts, but logically explores the unintended consequences as well. For example, using the futures wheel the logical implications of creating new roads in an undeveloped city may be explored. Economic activity may increase, leading to more jobs, higher prices, income rise, and more cars on the road. There may be more road traffic, pollution may increase and lead to health problems. The construction of roads may change the connectivity of social networks in previously more isolated areas. This may increase or decrease equity depending on rapid growth.

4.2 Specific Use of Futures Wheels for Interaction

The workshop for work package 5 featured the use of futures wheels as a vehicle for small group discussion and brainstorming. The workshop schedule, which was previously presented in section 2, is duplicated below for ease of discussion and reference. Futures wheels were used as part of a visioning and scenario creation exercise. Two 45 minute sessions were dedicated to the futures wheel activity. This is shown in the table below, under session 3 and session 4.

Table 7. Workshop agenda for Futures Wheels

Time			Topic	Mode
8:30	9:00	Registration and Introductions		
9:00	9:15	Agenda	Briefing	
9:15	10:45	Session 1. Review of Previous Findings	Moderated Discussion	A systematic review of the Big Data issues identified in seven sectors.
10:45	11:00	Coffee Break		
11:00	12:00	Session 2. Problem Structuring, Root	Moderated Discussion	Reviewing the critical assumptions and underpinnings of European governance.
12:00	13:00	Lunch		
13:00	13:50	Session 3. Vision/ Futures	Group Work	Specifying four future visions of European Big Data, with milestones for action.
13:50	14:40	Session 4. Scenario Creation	Group Work	Discussing actions to hedge and shape our European Big Data future.
14:40	15:00	Coffee Break		
15:00	15:45	Session 5. Keynote Soeech	Speech	Anticipating the need for privacy by design in Big Data systems and services.
15:45	16:00	Next Steps and Conclusions	Briefing	

Futures wheel templates were created and printed on A2 paper. Participants were provided with a list of previously identified external forces based upon the policy analysis templates. These impacts were written on post-its, and participants were invited to place the forces on the template as part of a group discussion. Three moderators aided and guided the discussion.

The figure below shows one such futures wheel, used as a basis for discussion by participants. In the center of the wheel a range of external forces are arrayed. New relevant forces, desccribed further in section 4.4, were introduced by participants. The outer segments of the wheel were used to initiate a discussion of possible policy levers for Europe. This topic is picked up for further discussion in deliverable D5.2.

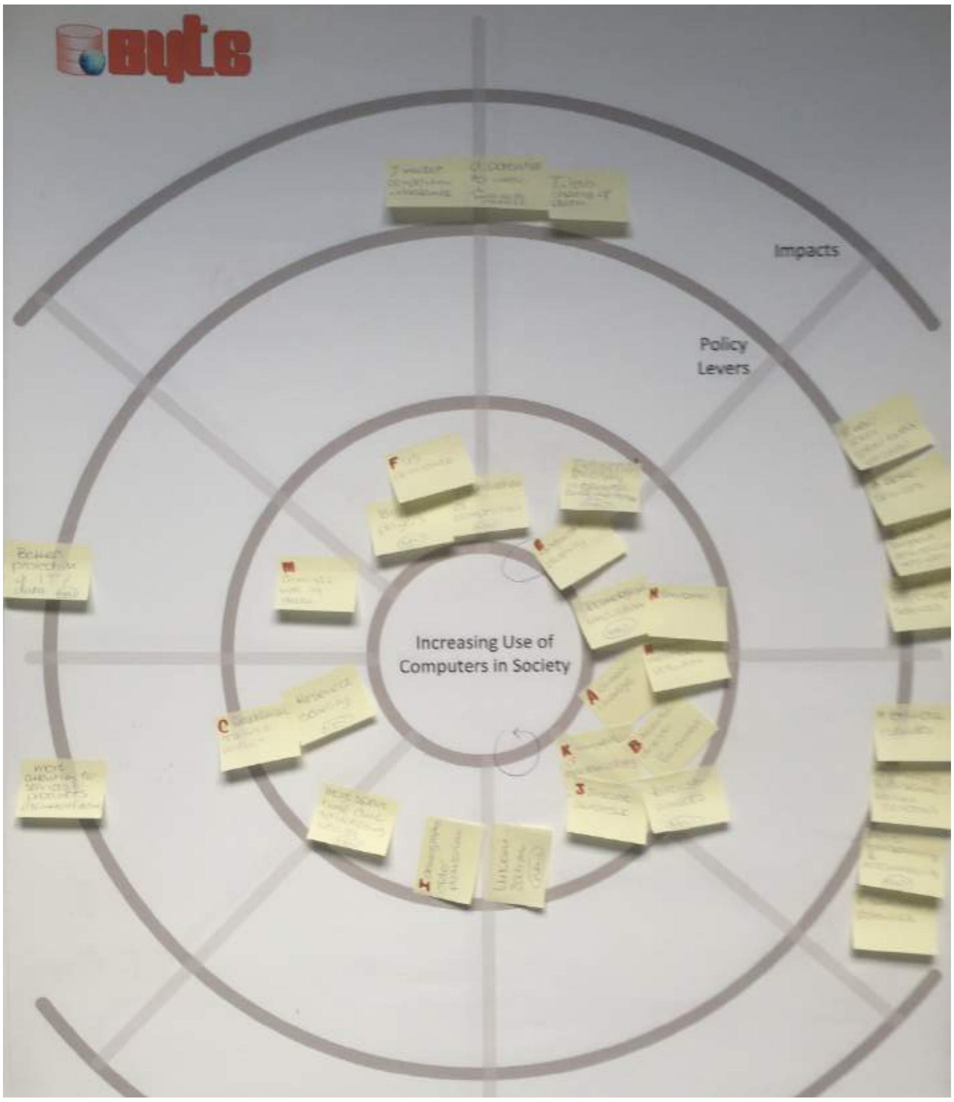


Figure 11. Futures wheel template

4.3 Specific Use of Futures Wheels as Input for Cross-Impact Analysis

As previously noted it is also possible to record the interactions of trend dyads using a qualitative cross-impact matrix. This is a square matrix with one row and column for each trend where the cells are filled with the impacts or effects of the trends. Both techniques generate a rich stock of material from which workshop participants can answer specific questions about potential futures, or even using the indicative impacts to write a complete scenario.

Conversely, but little explored, is the possibility of using futures wheels as input to a cross-impact analysis. The semantics of the futures wheels is that the radial placement of impact suggests the degree of relatedness between two forces. So, two forces placed coincident on the diagram receive a maximum rating of relatedness (for instance five on a five point scale). Two forces placed opposite each other on the wheel occur in disjoint futures, and are rated at the lowest possible rating of one on a five point scale. No meaning can be attached to the specific degree on the wheel, rather, it is the pattern of impacts and relationships in each specific wheel which is significant.

More specifically, during the workshop participants broke out into three separate groups. The groups generated three distinct futures wheels and discussed their results in a plenary session. Participants were also invited to add their own new external forces if and where they made sesnse. In addition, not all of the forces were selected or placed by all groups. This was a deliberate expression by the groups of which forces they believed would be most important for the big data future.

After the workshop the various forces on the diagram were encoded using a 360 degree schema. The corresponding cross-impact pattern was derived, and the cross-impact results were then added to the analysis results. Since each futures wheel is generated by multiple participants, the results were weighted proportionately more heavily by group.

A useful feature of the cross impact analysis procedure is that the procedure is indifferent to rotation. So if two groups had the same cross-impact pattern, but the pattern is rotated 120 degrees, the procedure recognizes the commonality in underlying forces. As previously noted in the cross-impact analysis section, the procedure enables a consensus weighting of the most significant forces, or combination of forces, as rated by users.

Participants seemed to enjoy the cross-impact exercise, and the task generally supported interactive discussion in a time-constrained workshop setting. There was some discussion and perhaps confusion regarding the diagram, particularly with regard to diametrically opposed forces. These forces were usefully interpreted by one participant as expression by a matter of degree. Alternative interpretations involved setting forces on or off, or understanding the polarity of the external forces. Some of the external forces provided were double sided, and therefore could potentially be broken up further into positive and negative expressions of external events.

One participant chafed at only having two major dimensions with which to place their forces and post-its. A compromise was brokered where otherwise independent forces were placed at the points of an equilateral triangle. This compromise results in a partial correlation or dependence between forces which was not necessarily intended.

4.4 SPECIFIC USE OF FUTURES WHEELS FOR BRAINSTORMING

The previous section (Cross-Impact Analysis) revealed three underlying trends in big data impacts. Furthermore this section argued that the forces as revealed by the various BYTE case studies were necessarily a partial, and incomplete, vision of the full trends surrounding the emergence of big data. Given this, it is appropriate to use creativity exercises to enrich and enumerate the variety of different forces which may underlying the transition to new regimes or hegemonies of big data.

In this subsection we therefore make further use of the futures wheel technique – as a brainstorming and creativity technique. In the brainstorming application of the futures wheel technique we describe new external forces as discusseed by workshop participants.In the creativity technique we describe possible new forces which emerge through a structured exploration of trends. The disccussion begins with the brainstorming exercise as conducted in the workshop.

Participants selectively introduced new external forces which may shape the European approach to big data.

N. Globalization
O. Factionalism
P. Terrorism
Q. Urbanization
R. Legislative Uncertainty
S. Systematic Technological Change

Figure 12. Additional external forces

Participants introduced a number of new social forces including factionalism – the social drive to create an in-crowd of like-minded people. This drive can be based on language, ethnicity or nationality. They also noted the possibility of major terrorist events. Globalism is a counter-acting force, at least partly economic, which enhances trade and communication but also redistributes the benefit of new technology. Urbanization, the increasing concentration of world population in cities, was noted.

One participant noted that big data might be part of a new an incipient technological revolution, much like the agricultural or manufacturing revolutions which came before in society. This perspective was distinctive enough from existing forces that it warranted being included. Note that while Europe is certainly an innovator on a world-scale, Europe can only partly steer or direct such change, making this potential revolution an external force rather than a policy lever.

Workshop participants also specifically called out legislative uncertainty as an important force worth considering. The diversity of legislative regimes concerning data across the world is a particular transaction cost for multi-national entities. Because these regimes are at least partly determined outside of Europe this is also a distinct external force worth consideration.

Next we turn to the use of futures wheels in an individual creativity exercise. The wheel shown in the figure below was the result of a creativity exercise where the trend variables of regime and hegemony are systematically varied, and a range of possible impacts and cross-impacts are listed. Moving outward in the wheel are the potential first, second and third-order effects identified through this exercise.

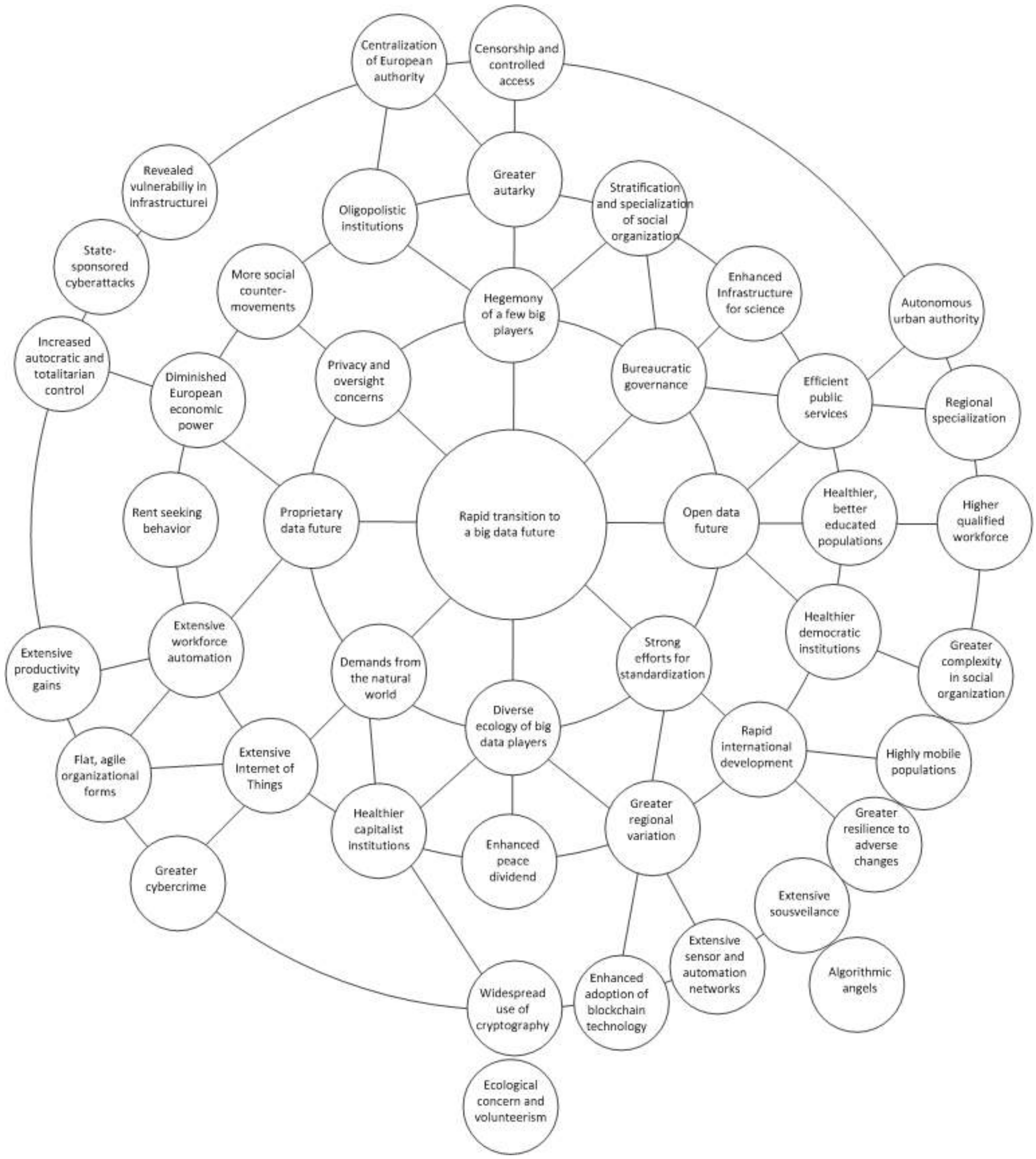


Figure 13. BYTE project futures wheel

The wheel shows a range of forces and impacts. Note that this wheel somewhat confounds the differences between external forces (which are not in European control) and impacts (which are the result of a willful selection of particular policies within Europe). Nonetheless this wheel shows the wide variety of potential societal and industrial segments which big data policies could potentially touch and reach. As an example, the choices of big data policies could affect the efficiency and affordability of public services, the health and education of European populations, and the speed of international development and welfare activities.

The external forces and impacts identified by the cases were somewhat circumscribed, perhaps given the sector specific and context specific character of individual cases. Nonetheless the cross-impact and workshop exercises have helped structure the kinds of impacts of most concern to BYTE experts and stakeholders. Naturally this sharper focus enables the project to examine a wider and richer set of possible forces and impacts.

The futures wheel serves as a useful precursor to scenario exercise exercises, particularly in the Manoa school of scenario studies. The following section outlines scenario analytic techniques in general, and in specific applies the identified forces and trends to create a small set of scenarios to support policy and planning.

5 SCENARIO ANALYSIS

Scenario development processes have exerted a strong influence on human thinking, decision-making process and public debate (Grunwald, 2011). They were primarily developed in the USA and France. Systematic use of scenarios to facilitate thinking about the future started after World War II. In the 1950s the US Department of Defense used it as a method for military planning at RAND Corporation (Kahn and Wiener, 1967; Joseph, 2000; Bradfield et al., 2005; Bezold, 2010). In the 1960s, scenarios were used extensively for social forecasting, public policy analysis and decision making. In the 1970s, scenario development in the industrial field was used as a strategic planning tool by the Royal Dutch/Shell Group. The technique was then popularized by Schwartz and Van der Heijden. In Europe, in the mid-1970s, Godet also began to develop scenarios for several institutions and companies, contributing to the development of the *La Prospective* school.

5.1 HISTORY AND PURPOSE OF SCENARIO ANALYSES

At a corporate level Shell is a widely known user of scenarios. They helped Shell cope with the oil shocks and other uncertain events of the 1970s (Joseph, 2000; Schoemaker and van der Heijden, 1992). Pierre Wack proposed three principles for scenario development based upon his experience in scenario planning at Shell: identifying predetermined elements in the environment, changing mindsets in order to perceive reality anew, and developing a macro level view of the business environment (Burt, 2010). Predetermined elements are historical actions and events that have already occurred or are likely to occur. It is critical to explore the consequences of these events because they act as the driving forces pushing for outcomes (Wack, 1985a;b).

In the present era characterized by uncertainty, innovation and change, increasing emphasis is being placed on the use of scenario planning techniques because of its usefulness in times of uncertainty and complexity (Schoemaker, 1993). There is consensus in the scientific community that scenarios stimulate strategic thinking and help overcome cognitive limitations by creating multiple futures (Curry and Schultz, 2009).

Scenario development and planning is a valuable method tool that helps organizations to prepare for future possible eventualities (Hiltunen, 2009). Herman Kahn one of the founders of futures studies, defines scenario as: ‘‘a set of hypothetical events set in the future constructed to clarify a possible chain of causal events as well as their decision points’’ (Kahn and Wiener, 1967). Broadly, scenarios outline some aspects of future, just as a movie scenario refers to an outline of the plot of a theater play or movie (Joseph, 2000). Scenarios provide the description of future situation and the development or portrayal of the path that leads from the present into the future (Pillkahn, 2008; Bishop et al., 2007). It follows from the definition that scenarios describe some future situation and the course of events-decisions will make this happen (Godet, 2000a). Scenarios are also defined as alternative futures resulting from a combination of trends and policies (Fontela and Hingel, 1993).

What needs to be emphasized is that scenarios are not a prediction of the future, but a record of the exploration of multiple plausible future situations. The objective is to help extend the thinking of scenario development participants (Godet, 2000b). They are a good way to question the future (Barber, 2009). Scenarios differ from forecasts because they contain a range of possible outcomes integrating uncertainty about the future whereas the purpose of forecasts is to identify the most likely pathway and estimate uncertainties (Pillkahn, 2008). Major benefits of the scenario approach include (Varum and Mel, 2010): improvement of the learning and the decision-making process, and the identification of new issues and problems which an organization or community may have to face in the future.

Exploration of many possible futures enables planning about the future in decision making processes (Burt and van der Heijden, 2003), and provides some margin to cope with uncertainty (Varum and Melo, 2010; Hiltunen, 2009). Future uncertainty increases as we move away from the present and look further into the future (Pillkahn, 2008). Scenarios provide an overall picture of significant trends that interact and are likely to generate events in the future (Martino, 2003). They present this in a coherent, systematic, comprehensive and plausible manner (Joseph, 2000). Through them it is possible to explore the implications of a particular choice or policy action, in light of possible future system discontinuities, their nature and timing (Strauss and Radnor, 2004). There are no guidelines as to the time horizon of scenarios but usually they are developed for years or decades (Martelli, 2001).

Perhaps the most common confusion when discussing scenarios is equating scenario development with scenario planning. The first involves thinking about the future and the uncertainty that surrounds it. Scenario development focuses more on creating actual future stories. Scenario planning is a more complete study of the future in the sense that it outlines possible futures and a range of action options (Schoemaker, 1991; 1993; Bishop et al., 2007). It follows that scenario development is the necessary foundation for scenario planning. Scenario planning is a way of facilitating adaptation to future major changes (Varum and Melo, 2010; Martelli, 2001).

Another clarification of terms involves the difference between scenarios and alternative futures. Defining scenarios narrowly would include only stories about alternative futures. This implies that other forecasting methods can produce alternative futures but not scenarios. In practice though the two terms are taken to be overlapping and equivalent. Little consideration is given to the creation of scenario stories in most methods and more is given to producing the core scenario logic which can be done by any number of methods (Bishop et al., 2007).

Scenario exercises or studies help to see the present under a different light, i.e. they are a devise for disturbing the present (Curry, 2009). This implies that ideally different options for developing and analyzing scenarios should be considered beyond the ones that fall within the operational and conceptual “comfort zone” of the organization (de Brabandere and Iny, 2010). This could result in exploring new possibilities and unique insights as different scenario generation methods yield not only different narratives and insights, but qualitatively different participant experiences (Curry and Schultz, 2009).

5.2 OVERVIEW OF SCENARIO APPROACHES

This section then provides a cursory overview of approaches to scenario development with some discussion of their strengths and weaknesses. A literature review reveals that there are several methodologies for generating scenarios with many common characteristics (Joseph, 2000; Chermack et al., 2001; Bradfield et al., 2005; Keough and Shanahan, 2008; Bishop et al., 2007;Varum and Melo, 2010). There are several methodologies and guidelines proposed in the literature for scenario development. Three review articles in the field of scenario development (Van Notten et al., 2003; Bradfield et al., 2005; Borjeson et al., 2006) are discussed next to provide the context and rational for some commonly used methods. Interestingly most techniques do not use computers in scenario development. It is perhaps an area of future opportunity to make greater use of software in crafting scenarios (Bishop et al., 2007).

Van Notten et al. (2003) propose a typology of "scenario types" with three major categories based on the scenario goals (exploration vs decision support), the scenario development process (intuitive vs formal) and the attributes of scenarios (complex vs simple). These categories constitute more overall scenario project characteristics than specific scenario technique(s).

Explorative scenario development may include awareness raising, stimulating creative thinking, and generating insights on how societal processes influence one another. The main aim is to go through the process and this is just as important as the end outcome of it. Scenario development used for decision support, is used to examine plausible and possible paths to futures that vary according to their desirability. It is possible that scenario development may produce concrete strategic options for each future. Scenario development for decision support often produces value-laden combinations of scenarios that can be: (i) preferable, optimistic, utopic, (ii) conventional, or middle-of-the-road, and (iii) disagreeable, pessimistic, dystopic, or doom scenarios.

The scenario development process categories relate to the degree of quantitative and qualitative data used, the choice for stakeholder workshops, expert interviews, or desk research. Completely informal or intuitive approaches rely strongly on qualitative knowledge and insights when developing scenarios. Techniques such as developing storylines in interactive group sessions with a high variety of people are typical intuitive approaches to scenarios (Schwartz, 1991). At the other end of the spectrum are formal approaches such as the *La Prospective* where scenario development is seen as a rational and analytical exercise. Formal approaches utilize quantified knowledge and often use computer simulation techniques. Finally, there are examples of scenarios that combine intuitive and formal process designs such as the IPCC emission scenarios (IPCC, 2000).

Finally, scenario attributes describe the nature of variables and dynamics involved in a scenario, and how they interconnect. Variables include actors, factors, and sectors (Rotmans, 2000). Actors can be individuals, organisations or groups of organisations such as governmental bodies, companies, NGOs and scientists. Factors are societal themes such as equity, employment, consumption behaviour, and environmental degradation. Sectors are arenas in society where factors and actors interact. Van Asselt (2000) argues that a decision-making process is complex when the following conditions apply:

- There an ensemble of connected problems.
- The issue at hand cuts across various scientific disciplines.
- Actors, factors and sectors interact on various scale levels.

It follows that a complex scenario is composed of causally related variables and dynamics which can manifest alternative patterns of development. Complex scenarios often draw on a broad range of actors, factors, and sectors, across time, space and scale. In contrast, simple scenarios are more limited in scope and may simply consist in extrapolating existing trends such as the European Environment Agency's baseline scenario on the future of Europe's environment (EEA, 1999).

In contrast to Van Notten's taxonomy that proposes attributes of scenarios, Bradfield et al., (2005) propose actual high level categories. Their approach is historical, tracing the evolution of three schools of scenario development from their origins to the present day. Two of these schools originate in US and UK and one in France. The first is the "intuitive logic" school or the Shell/GBN method that now dominates scenario development in the USA and many other countries. The second is the "probabilistic modified trends" school by Olaf Helmer and Ted Gordon. This technique is quantitative, as opposed to the Shell/GBN technique.

Borjeson et al. (2006) create a typology of scenario techniques based on different types of probable, possible and preferable futures. Predictive scenarios answer the question: "What will happen?" Exploratory scenarios answer: "What can happen?" Normative scenarios answer: "How can a specific target be reached?" They divide each of these into two sub-categories to make six types of scenarios (Figure 1).

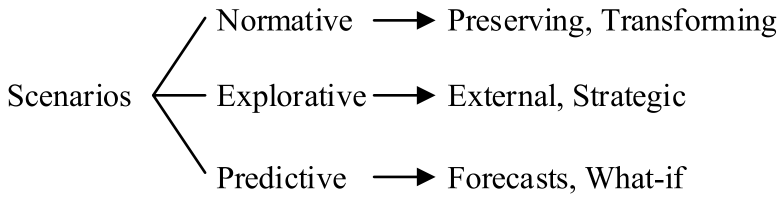


Table 8. Categories of scenario techniques (adapted from Borjeson et al., 2006)

The process of scenario development includes a number of tasks and related techniques according to Borjeson et al. (2006). First, comes data gathering and generating ideas. Generating techniques are techniques for generating and collecting ideas, knowledge and views regarding some part of the future, consisting of common data gathering techniques such as workshops, surveys, and interviews. A widely used technique is the Delphi method and its variants (Borjeson et al., 2006) where the aim is to collect and harmonize the opinions of a panel of experts on the issue at stake.

Second, is the integration of parts into wholes. Integrating techniques are frequently based on mathematical modelling and use models based on quantitative assessments of probability or relations such as time series analysis and systems models. The end result of these techniques is future projections under explicit constraints frequently quantifying uncertainty. However, the quantification of uncertainty often depends on subjective assessments of the likeliness of various events.

Third, the consistency of scenarios needs to be checked. Consistency techniques include morphological analysis and cross-impact analysis that ensure consistency among different forecasts. In practice, consistency testing is often carried out in a qualitative and sometimes implicit way, e.g. by using expert panels to get critique and suggestions for improvement. Finally, Bishop et al. (2007) identify eight categories of scenario techniques. From these, the

Manoa and FAR techniques are discussed in more detail later in the document as they are applied in work package 5.

1. Judgment (genius forecasting, visualization, role playing, Coates and Jarratt)
2. Baseline/expected (trend extrapolation, Manoa, systems scenarios, trend impact analysis)
3. Elaboration of fixed scenarios (incasting, SRI)
4. Event sequences (probability trees, sociovision, divergence mapping)
5. Backcasting (horizon mission methodology, Impact of Future Technologies, future mapping)
6. Dimensions of uncertainty (morphological analysis, field anomaly relaxation, GBN, MORPHOL, OS/SE)
7. Cross-impact analysis (SMIC PROF-EXPERT, IFS)
8. Modeling (trend impact analysis, sensitivity analysis, dynamic scenarios)

Bishop et al. (2007) identify two variations on category 2: one that elaborates the baseline scenario using futures techniques and one that adjusts it given the occurrence of potential future events. The first one is the Manoa technique, developed by invented by Wendy Schultz and Jim Dator at the University of Hawaii. It is an amalgam of futures techniques to explore the implications and interconnections among trends.

A variation of the Manoa approach called Systemic Scenarios was developed by two of Dr Schultz's students, Sandra Burchsted and Christian Crews, also developed a variation of the Manoa technique that they call Systemic Scenarios (Burchsted and Crews, 2003). Rather than use a cross-impact matrix to map interactions among trends, they use a causal model which shows the dynamic interactions among the implications and hence the trends (Burchsted and Crews, 2003).

5.3 Specific Foresight Methodologies

5.3.1 Manoa School Scenario Approach

This approach was designed to maximize difference from the present, as a means to challenge it. The process, focusses on helping people understand the dynamics of change rippling through various systems, as drivers create primary, secondary, and further cascades of impacts, which then create cross-impact turbulence (Curry and Schultz, 2009). The Manoa approach assumes that actual futures are generated by the turbulent intersection of multiple trends, and the interplay of their cascading impacts (Curry and Schultz, 2009).

The scenario development approaches presented by Dator and others, are depth based, rather than breadth based and instead of dealing with uncertainty, these scenarios expose power relations. As Dator claims: ''What futurists can and often do study, are 'images of the future' in people's minds'' (Dator, 2002). He further elaborates that future studies often serve as the basis of actions in the present. Dator's work on his alternative futures approach articulates four scenario archetypes (Dator, 1979): continuation, collapse, disciplinary society and transformational society through high technology or high spiritual change. Taking these four scenarios, one can envisage how future would look in each of these scenarios.

Developed by James Dator, the Manoa method articulates four scenario archetypes (Dator, 1979; www.futures.hawaii.edu). Taking these four scenarios, one can incast or articulate how one's organization would look in each of these scenarios. None of these four futures is intended to be any better, or any worse, than any other. They can be positive or negative to stakeholders that have their vested interests influenced somehow, and they should be presented positively. The scenarios are:

- *Continued growth:* In this future, it is assumed that current conditions and trends are enhanced. This implies that there are more products, more infrastructure (e.g. roads), more technology, and a greater population. It is assumed that technology is the solution to every problem.
- *Collapse:* This future comes about because of continued growth failing. Tensions inherent in societal systems are too great to be borne and lead to collapse: in socioecological system interfaces, in gender equality, in financial/economic institutions, cultural-religious, and sociotechnical.
- *Steady state:* This is a future where a balance is sought between the economy and nature and seeks to arrest growth. A more balanced, equitable, community driven society is the outcome. A steady state scenario involves a return to nature and human values where technology is often seen as part of the problem.
- *Transformation:* This is a future where the core assumptions of the previous three are changed. Thus, transformation involves either a dramatic technological change of broader societal repercussions (artificial intelligence eliminates the courts, bureaucracy and many forms of governance; genetics changing the nature of nature) or through spiritual change (humans change their consciousness, not just values, but the experience of deep transcendence).

Assumptions underlying the four generic alternative futures Dator (2009)

Rationale for Continued Growth Scenario: This is the most commonly assumed of the four alternative futures since almost all official statements about the future are based on continued growth. A state of continued growth is official dogma in most modern governments, educational systems, and organizations. Broadly it conveys the persistence of the general characteristics of American society: growth oriented, opportunity-filled, technologically-progressive, upwardly-mobile, internationally-dominant, science-guided, rich, leisure-filled, abundant, and liberal. Under this rationale, it is the purpose of government and education to build a vibrant economy, and develop the people skills, institutions, and technologies to keep the economy growing endlessly.

Rationale for Collapse Scenario: This reflects the concerns of people about social and/or environmental collapse according to which the economy cannot and possibly should not be allowed to growing endlessly in a finite world. There is a range of reasons why people may anticipate some sort of collapse: economic, environmental, resource, moral, ideological, a failure of will to change. Alternative plausible sources of collapse involve invasions of countries, or cataclysmic events (tsunamis, earthquakes, global warming, pandemics, space meteorites etc). The administrative inability to deal with a snowballing series of pure accidents or terroristic events, nuclear war or some combination of these may lead to a trajectory of collapse.

Such sort of events or any combination of them, may stress our limited, fragile and globally interconnected and interdependent world to collapse leading to a sudden population collapse or a dystopian new Dark Ages. This kind of scenario does not necessarily apply on a global scale only. Communities, organizations, and cultures face economic and social forces that may render once-valuable institutions obsolete or unviable. It is also possible that communities, or even countries destabilize and sink into a state of perpetual disorder. The collapse scenario is not the same as a worst case scenario as there may winners and losers after such an outcome.

Rationale for Steady Stable Scenario: Steady state is sometimes used as a synonym for disciplined society scenario. This scenario appeals to people's concern that the effects of continued economic growth are either undesirable or unsustainable. While continued economic growth might be good or necessary to alleviate the large disparities evident in the world today, it also becomes realized that continued economic growth is unsustainable on a finite planet where resources are rapidly depleted from a growing population.

Modern technology has provided the capacity to attend to an ever increasing array of human needs for an ever increasing population, even going beyond the natural capacity of the planet's resources. Nevertheless, a state of continued growth may be slowly but inexorably coming to an end along with cheap and easily available energy resources, and with an increasing waste production by humans. What is needed is a managed downscaling rather than growth.

Against this dire future, it is argued that it is necessary to refocus our economy and society on survival and fair distribution, and not on continued economic growth. A steady state scenario allows the preservation and/or restoration of places, principles, and processes that humans find important and attach higher value to, compared to materialist gains. This implies reorienting personal lives as well around a set of fundamental values: natural, spiritual, religious, political, or cultural and finding a deeper purpose in life than the pursuit of endless wealth and consumerism.

Rationale for Transformation Scenario: The transformation scenario reflects a future where the power of technology especially robotics and artificial intelligence, genetic engineering, nanotechnology, teleportation, space settlement, overshadows the effects of other restraining trends and brings about the emergence of a transformation society as the successor to the information society. In some sense this scenario parallels the steady state scenario but foresees new values, institutional, and technological arrangements fundamentally different from anything seen before. It is possible that transformation will be driven by fragmenting and individualizing effects of novel technologies, and the end to Western dominance. The scenario anticipates and welcomes the transformation of all life, including humanity from its present form into a new "posthuman" form, on an entirely artificial Earth, as part of the extension of intelligent life from Earth into the solar system and eventually beyond.

5.3.2 Field Anomaly Relaxation

The Field Anomaly Relaxation (FAR) method was designed to address the task of generating a set of representative scenarios which are reasonably exhaustive for a domain of interest (Rhyne (1974;1981; Cappert, 1973; Wood and Christakis, 1984). FAR has been applied to a wide range of domains including military planning, educational provision and water resources management. At its core FAR deals with the task by utilizing several variables or "factors" to describe the present and future states of a domain of interest. This requires wide-ranging thought about the future.

FAR uses several factors that jointly define and describe a domain of interest and a partiular socioeconomic context. For example, relevant factors can be the level of prosperity, the degree of social cohesion, and the stability solidity of international agreements. Each of the factors is further elaborated in terms of the possible levels of intensity or levels of manifestation that have occurred in the past and can occur in the future (Figure 3).

When a number of factors are combined they produce a field. It is likely that some combinations will not be logically plausible, for example excessively high levels of terrorism cannot coexist with high prosperity (Rhyne ,1981). In FAR terms this is an "anomaly:" a future state that cannot be reached. A range of such future states may exist both at final and intermediate points in time. However a time line that avoided any anomalies would represent a possible future. Thus, even in a field involving four or five variables, each defined at five or six levels, which could involve many tens of thousands of theoretical combinations, the elimination, or 'relaxation', of anomalies would reduce the number of possibilities to a few score and the number of time lines to a handful.

FAR does not claim to derive the most probable or most plausible futures, and it does not discuss probabilities. It is intended to bring scattered information and insights together, and facilitate the composition of a smaller set of alternatives from which internal inconsistencies have been removed. Each future projection gives a set of logically plausible scenarios for a selected domain. This yields a set of comparably plausible scenarios for a chosen field, each describing changing circumstances rather than events over a future span of one to three decades.

The strength of FAR is its ability to reject scenarios that are logically inconsistent in a traceable manner, thus equally providing information from less or more likely scenarios. FAR uses trend extrapolations only tangentially, as often to help expose unlikelihood as to predict the future. Its steps are iterative and ordered, but mostly qualitative. A FAR exercise consists of four steps, which may be repeated more than once if necessary. In this case the results from the first iteration will serve as the inputs for the second. The FAR steps are as follows (Gappert, 1973; Rhyne, 1974):

Step 1. This steps starts with imagining alternative futures for the domain of concern. One is selected as an initial visualization of the future without it being the most plausible necessarily. This choice may be done in the first of a series of FAR cycles. Alternatively a complete description may generated using other sources.

This initial future description serves as a basis for the selection of factors that describe the future state in its entirety. Then the levels of the factors need to be determined. There are two rules for doing so: (i) the factors must be as mutually exclusive as possible, and (ii) they must represent a range of plausibility. The levels should be selected only on the basis of factor states that could exist. No consideration is given at this stage to how these states will come about, although one of the factor descriptions for each sector must describe the current situation. Then the choice of levels should: (i) define the two extremes that may possibly occur, (ii) define the current situation, and (iii) be sequential in change definition i.e. if a factor is at a certain level it can only change to the one above or below, it cannot skip a factor.

Step 2. This step involves grouping the factors into about six sectors, which summarize the multiple descriptors and are relevant to the chosen domain. They should be chosen without assuming a relative dominance among them.

Step 3. This step involves "relaxing" or taking out the inconsistent future configurations in the array constructed in step 2. If any factor pair is too implausible to merit further consideration, the internal consistency of the entire future projection can be flawed, and should be eliminated. This step needs to be repeated as there may be a lot of futures surviving the pair-wise comparison, and because not all plausible pairs in a given future will form a coherent whole when brought together.

The end results of anomaly relaxation is a number of logically consistent plausible future visualizations. This process should: (i) consider sector-to-sector comparisons alone in matrix form (but other matrices should be available for reference) (ii) define a level of unanimity among participants for identifying anomalies, (iii) attribute comparative possibilities to more or less plausible scenarios generated as a result of anomaly relaxation.

Step 4. In this step, the remaining configurations are combined into distinct plausible sequences to produce scenario outlines. Several of the more plausible scenarios are chosen and written in narrative form to give an understandable description of the particular scenario. This could serve as the first step in second iteration. If some of the futures have common elements in their fields then this increases confidence in them. Nevertheless, it would not provide ground with which to treat the results as predictions as even overlapping fields are just the best understanding currently available about how the future might unfold. Linking configurations together to create plausible time lines is done in four distinct stages: grouping, merging, creation of time lines, and scenario generation.

Grouping. This involves deciding how soon configurations can happen and clustering them in groups. The groups cover successive five-year steps into the future. Carrying out the grouping of configurations involves subjective judgment.

Merging. Related configurations are merged using the distinctively different criterion. Two simple rules are applied in this step: (i) if the only difference between a grouping of configurations is variation in sectors that is less important than the remaining sectors the configurations are merged, (ii) if all the configurations in a particular grouping have identical levels for the majority of their sectors, then they are merged together despite the difference in other sectors levels.

Creation of time lines. This step involves selecting possible time lines by creating plausible links between groups of configurations.

Scenario generation. The final step is to validate the procedure through a separate attempt at generating a narrative version for a chosen future scenario and comparing it to the output of the FAR process to see whether the results are similar enough to be valid.

5.4 Scenarios in the BYTE Vision

The following section outlines how scenarios are derived from the cross-impact analysis and futures wheel methodologies as described in sections 3 and 4. Many scenario building approaches use a qualitative logic approach. The trends identified serve as a qualitative basis or scaffold for building such scenarios.

The simplest possible structure for expressing three distinct trends requires a 2 x 2 x 2 strategic framework. Eight scenarios result. The first dimension would express high and low rates of technological transition. The second dimension expresses two styles of hegemony. The third dimension expresses two regimes of big data. As a result eight distinct futures can be described as the combinations of these three forces.

The BYTE vision consists of two sets of four scenarios, distinguished along three major trends. As discussed previously the trends are the big data transition. This indicates how rapidly the world progresses towards big data, and the technologies beyond. Will the transition occur rapidly, or will it be a tepid or stalled transition? Another trend which constitutes these scenarios is the big data hegemony – will the big data ecology be controlled by a handful of big players in the public or private realm, or will it be characterized by a diverse marketplace comprised of many players? The final trend is the big data regime – will big data be governed in an open fashion, like a common pool resource, or will it be governed in a closed fashion like intellectual property?

We can also collapse or simplify this scenario logic. It might be most productive to focus on the four scenarios surrounding a rapid transition to big data. That is, we may wish to drop off the scenarios involving a stalled big data future in order to focus on those scenarios with the greatest amount of technological change. These rapid change scenarios are shown below, focusing more specifically on the qualitative logic of hegemony and regime. It is not required that the scenario labels follow the general Manoa school logic of steady state, continued growth, transformation and collapse. Nonetheless it is instructive to see the Manoa school scenarios overlaid on the BYTE trend logic. The Manoa names and narratives have been seen and reproduced across a wide variety of potential cases and domains.

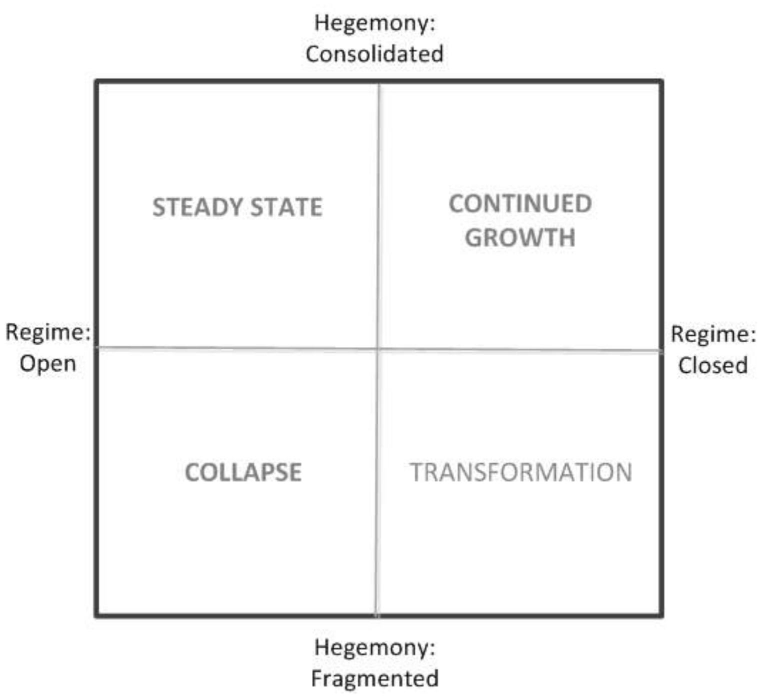


Figure 14. Manoa school scenarios and logic

The previous cross-impact methodology enables us to embed the seven BYTE cases within this scenario logic framework. This occurs because a unique set of external impacts were identified in each of the BYTE cases. Although each of the external forces are unique, participants see and anticipate underlying causes and relationships between these forces. Thus, it is possible to put each of the cases in the location where the fullest expression of underlying forces is seen. This is demonstrated in the figure below, and more fully discussed later in this section.

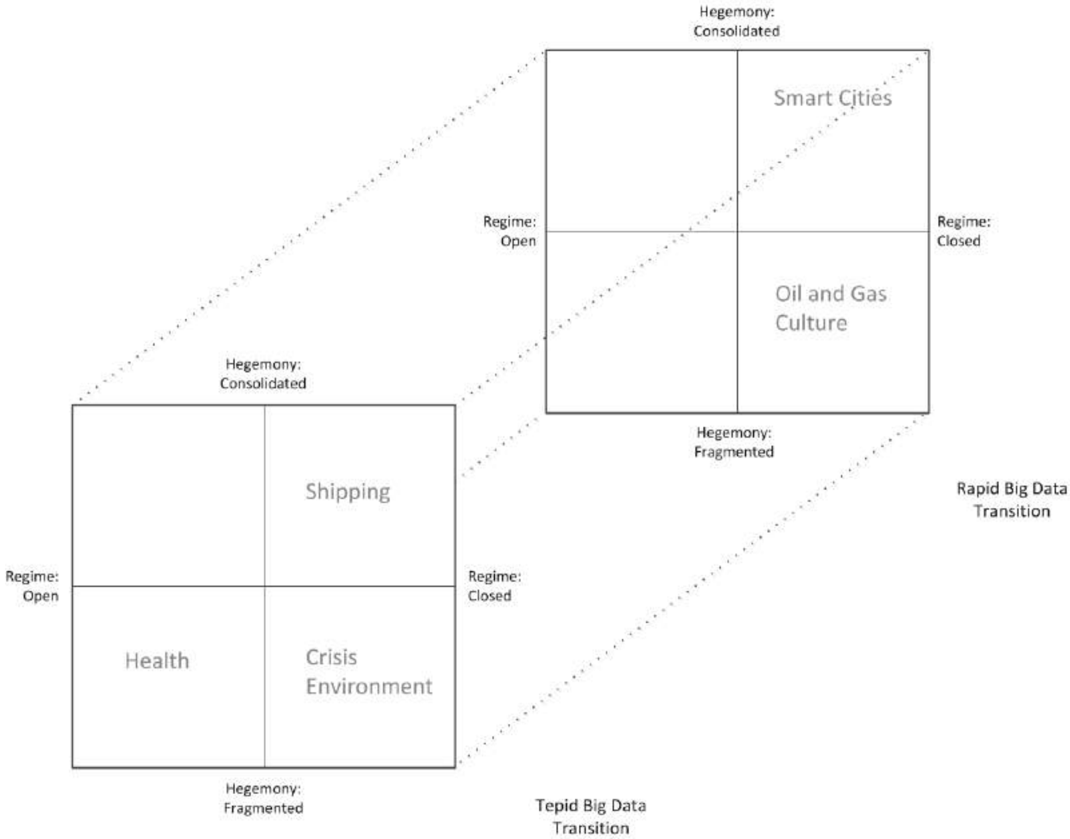


Figure 15. BYTE scenarios and cases

Qualifications are in order. Some of these trends could usefully be further explored or diversified. For instance a big company hegemony is likely to be very different than a big brother kind of hegemony. Each of the forces is also a matter of degree. Any given big data regime is likely to exist on a continuum from open or closed policies. There may even be different regimes in different industrial or economic sectors..

In fact these four scenarios are explored more thoroughly in four narrative vignetes, in the figure below. These four cases are given from the perspective of local government – the Delft municipality, a medium-sized city in the western part of the Netherlands. The goal of these short narratives is to provide a vivid portrayal of how very different four big data futures may actually be. Other possible stakeholders, and the values and consequences for stakeholders, is discussed in deliverable D5.2.

The local government of Delft has increased its tax base and minimized unnecessary expenditure through the use of big data. Smart metering systems enable the local government to closely track local water and waste statistics, and to impose a highly efficient set of tariffs on users. Routine city maintenance tasks have been automated, and are steadily improving through machine learning systems. Municipally owned combined heat and power systems constitute a small yet growing source of revenue for the city. A media wall in the city hall provides real-time tracking of civic participation and satisfaction.

Policing is greatly assisted by a fleet of drones which are used to track suspected arsonists and burglars. Beat patrols are given new routes in response to crime statistics and to the sentiment analyses of local media. Close-circuit video cameras are spread throughout the city centre; facial recognition software records the movement of specific targets of interest. Libertarian activists raise new concerns about police targeting of minorities. Students and intellectual migrant workers argue that the regime of taxation and fines is mercenary, particularly for those who do not have a long-term residency permit in the city. There is a significant period of turbulence as appropriate legislation is developed.

Figure 16. Continued growth vignette

The city and citizens of Delft have negotiated and accepted a new balance between pubic and private interests in big data. A number of distinct arenas for big data governance have been created, ranging in their degree of openness. Appropriate public safe-guards are made to ensure the privacy of citizens. This requires a three-way surety between member states and the central European government. A radical system of "surveillance from within" has emerged in Europe, enabling the municipality of Delft -- when in need -- to track its citizens across both physical and virtual time lines. Europe has benefited from a harvest of open data, with corresponding gains in cooperation and coordination across a range of environmental, public health, and infrastructural services.

Many companies are unable to appropriate a sustainable advantage by collecting their own unique sources of data. Small and medium-sized enterprises find themselves unable to compete in European-wide data initiatives. Innovation has been stifled by the bureaucratic oversight required for these systems. Fundamental legislative uncertainties remain as various localities in Europe exercise their prerogatives for data protection and forgetting.

Figure 17. Steady state vignette

The local government has signed a one hundred year lease with a world-renowned civic technology company. A sensor-network throughout the city continuously captures and stores a range of data; the city gains free storage and back-up for all the data. The data is compared and benchmarked with peer cities across the globe. An administrative dashboard, containing the most actionable metrics, is accessible for city officials using their smart phones. A 24-hour professional services and response crew is on call to deal with any emerging crises. This crew is able to develop services on demand to meet the unique needs of the Delft municipality.

The actual data of the city is owned by the civic technology company. Concerns have been raised about whether local capacities in governance are maintained. The most talented civic technologists, who can audit and customize the software of the multinational, are in great demand through-out the world. Unfortunately the local municipality is unable and unwilling to compete for these professionals. The city faces a hold-up problem as the civic technology company upgrades to the next generation system. Local activists are pressuring the municipality to track farming specific emissions in water and air. Unfortunately these metrics are not tracked.

Figure 18. Transformation vignette

A new ambient intelligent system took over parts of Delft and the Westland. These systems, known generically as genius loci, are increasingly common in heavily urbanized and networked locations throughout Europe. The system coalesced out of the internet of things and other distributed contract management systems. Delft, like other cities of its kind, receives a range of useful services from the system including anonymization, counter-intrusion, indirection, localization, relationship management, preventative maintenance and relationship management services. All the city information is for free, suitably anonymized, for anyone knowledgeable to access it.

Older ceremonial and political boundaries have been steadily eroded by the genius locii. Zoning rules and regulations have been supplanted and sides-stepped. This trend dismays many political actors which never tacitly accepted the help of the genii. The price of anarchy has also proven to be quite high with these systems. Minor boundary disputes with the Haagland genius often escalate, sending fleets of automated vehicles throughout the narrow Delft streets. In many ways it has been easier for the new emerging megacities of the world to deploy these systems.

Figure 19. Collapse vignette

5.5 Sector Specific Visions

The final part of this section are sector specific visions. Our cross-impact work indicates that four of the BYTE case studies are anticipating a slow or tepid transition. These include the health, shipping, crisis and environment cases. Stakeholders in these cases are naturally more concerned with the real issues happening in humanitarian aid, health care and the environment. Nonetheless these sectors may be ill-prepared should a rapid transition to the big data world occurs, where the virtual and electronic worlds play a much bigger role in determining the impact and success of endeavours in the natural world. On the contrary, the smart cities case, the oil and gas case, and the culture case are already living in, and anticipating, a range of big data futures.

Another of the trends underlying the vision is the hegemony of big data. The cases are comparatively split on this trend. The smart cities and shipping cases seem well-prepared for a future with a few big players in big data. In contrast, the health, crisis, environment, oil and case and culture cases are all anticipating working with a diverse set of actors. Here again there are two major concerns – which futures are more likely, and which futures are most desirable for Europe. The role of this vision is to open up discussion to a variety of possible uncertain futures, rather than to predict or recommend just one.

As noted, one of the major trends in the BYTE vision is the regime of big data governance. Fully six of the seven cases anticipate governance in a close regime – shipping, crisis, environment, smart cities, oil and gas, and culture. Only the health care experts and stakeholders seem prepared for a very different future where data is openly and freely shared. The BYTE sectors may be woefully unprepared. The open data futures may deliver more desirable outcomes overall for these sectors. Or the open future may be coming anyhow, and these sectors will be caught without adequate preparation or governance.

One means of evaluating the BYTE case study work is in terms of its generality. Part of the generality of these cases depends on whether a complete, diverse set of external forces have been identified. By this measure the BYTE case work has done a fair job of selecting relevant sectors. Seven cases occupy five of the eight possible futures as identified by the vision. More worrying though is that most of the cases have not fully imagined the scope and extent of possible technological change in a digital world. Even more concerning is the fact that too few of these cases anticipate an open world, particularly an open world which is governed by a few, consolidated authorities. Europe may experience such a world, or it may wish to create or take part in such a world.

6 RELEVANCE FOR OTHER SECTORS

6.1 POTENTIAL FULL RELEVANCE FOR BIG DATA

The cases in the BYTE project were selected with a particular logic. Cases might well have been chosen randomly; this kind of selection is neither necessary, nor preferable for the purposes of the project. The case selection might appear as a wrong approach if case selection is assumed to produce a group of cases representative of some general population of organizations or firms in the case of big data. The aim of the BYTE project was not to create such generalizations. The cases were selected because they were unique, they provided contrasting examples, and enabled elaboration on existing phenomena and trends.

The choice of particular organizational settings was also not random but purposive in the sense that it allowed particular insights that other organizational settings would not allow (Siggelkow, 2007). Case variety in search of unique insights was something that was purposively done. Moreover, multiple cases enable comparisons that clarify whether an emergent finding is simply idiosyncratic to a single case or consistently replicated by several cases (Eisenhardt, 1991). The emergent insights are more robust because they are grounded in varied empirical evidence. Each additional case study doubled the analytic power of the research undertaken (Eisenhardt and Graebner, 2007). In the context of the BYTE project, the choice of the case studies was made aiming to maximize the diversity of the cases in terms of the characteristics of big data: velocity, volume, and variety.

Even following the exact opposite logic to case variety in choosing the cases, there would still be industrial sectors left that the project wouldn't cover through cases. In either way then a broad, all-encompassing view on the impact of big data, requires an additional look into the remaining sectors of the economy. A step towards this was made by looking at the standard industrial code classification (SIC) that United Nations produces (2008) listed in the table below.

Table 9 Standard Industrial Codes (SIC)

A. Agriculture, forestry and fishing
B. Mining and quarrying
C. Manufacturing
D. Electricity, gas, steam and air conditioning supply
E. Water supply; sewerage, waste management and remediation activities
F. Construction
G. Wholesale and retail trade; repair of motor vehicles and motorcycles
H. Transportation and storage
I. Accommodation and food service activities
J. Information and communication
K. Financial and insurance activities
L. Real estate activities
M. Professional, scientific and technical activities
N. Administrative and support service activities
O. Public administration and defense; compulsory social security
P. Education
Q. Human health and social work activities
R. Arts, entertainment and recreation
S. Other service activities
T. Activities of households as employers; undifferentiated goods and services-producing activities of households for own use
U. Activities of extraterritorial organizations and bodies

The BYTE cases were then classified under the SIC codes. There was no exact match, and some cases appear to be related to more than one SIC codes. The crisis case comes under the code for: information and communication (I), human health and social work activities (Q), and activities of extraterritorial organizations and bodies (U). The culture case comes under the code for arts, entertainment and recreation (R). The environment case comes under the code for water supply; sewerage, waste management and remediation activities (E). The health case comes under the code for human health and social work activities (Q). The oil & gas case comes under the code for mining and quarrying (B), and transportation and storage (H). The shipping case comes under the code for transportation and storage (H). The smart cities case covers aspects coming under the code for: information and communication (J), administrative and support service activities (N), and public administration and defense; compulsory social security (O). The classification is summarized below.

Table 10 Classification of cases under SIC

Case	SIC
Crisis	J, Q, U
Culture	R
Environment	E
Health	Q
Oil & gas	B, H
Shipping	H
Smart cities	J, N,O

The remaining sectors that do not overlap with the cases of the project are listed below:

A. Agriculture, forestry and fishing
C. Manufacturing
D. Electricity, gas, steam and air conditioning supply
F. Construction
G. Wholesale and retail trade; repair of motor vehicles and motorcycles
I. Accommodation and food service activities
K. Financial and insurance activities
L. Real estate activities
M. Professional, scientific and technical activities
P. Education
S. Other service activities
T. Activities of households as employers; undifferentiated goods and services-producing activities of households for own use

6.2 Grouping of sectors by externalities

Many of the additional sectors are likely to be influenced by the external forces already identified in BYTE cases. Because these sectors face similar external forces, they are also likely to benefit from similar policy levers. Furthermore, even if this is not the case, policy responses to external forces documented in the BYTE cases could have a knock on effect to the additional SIC codes listed above. Therefore it is worth exploring this kind of synergy between external forces and sectors listed above.

6.3 Judgmental approach for scoring new sectors

The list of remaining SIC codes is paired with the list of external forces identified in the BYTE project. The aim is to identify which external forces cut across them. Each external force can affect an SIC code with an impact of 1 (low) to 5 (high). This then used to complete a table. The process has been repeated 3 times and results are presented below

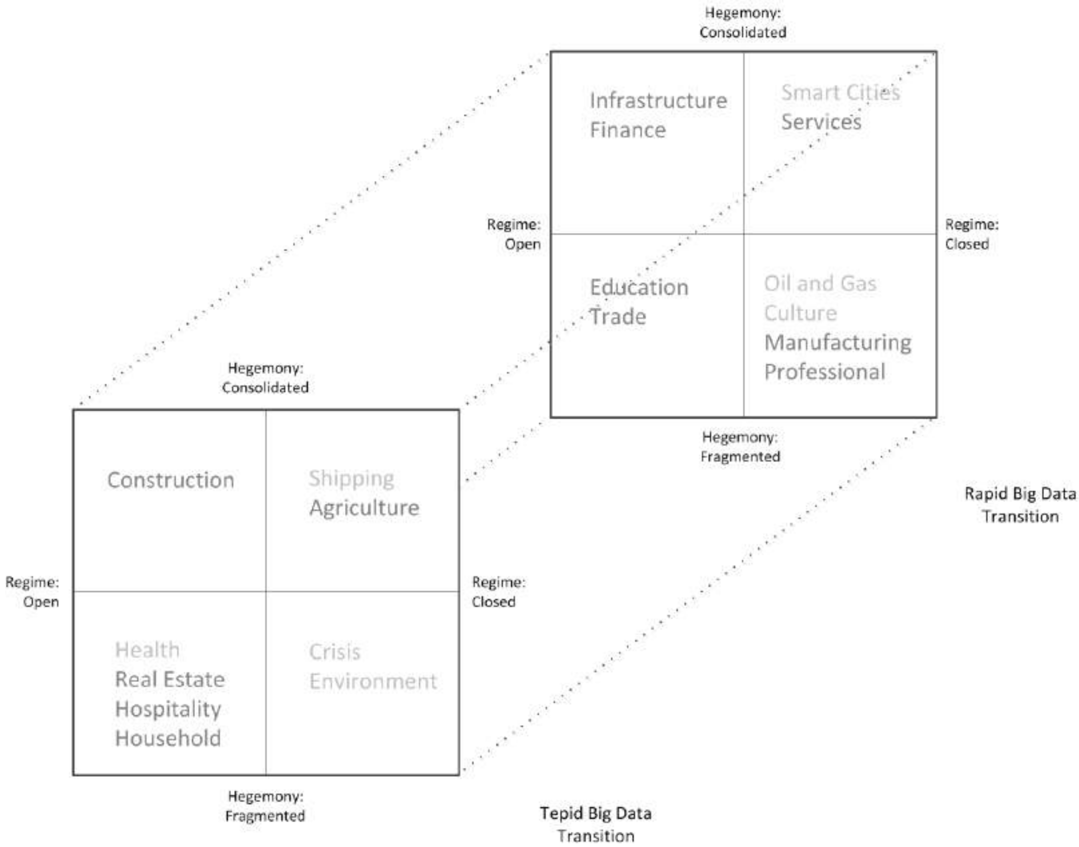


Figure 20. Further relevant sectors

As discussed in the previous section we evaluated twelve additional industrial sectors using a consolidated list of BYTE external forces. This permits us to evaluate and generalize the sufficiency of the existing BYTE cases, as well as anticipate possible new and emerging externalities. These twelve new sectors round out and fully occupy all eight of the scenarios in the BYTE vision. The BYTE cases are shown in a lighter font. The other twelve industrial sectors, providing a complete view of the economy, is shown in a darker text.

Again it should be emphasized that not all of these twelve scenarios in the vision will necessarily come to pass. Perhaps only one of these scenarios will come to pass. Or perhaps several of these scenarios come to pass, as different industrial sectors evolve in their own specific manner. It is the purpose of the BYTE project to ensure that a comprehensive scanning exercise, across all sectors of the economy, has occurred. The vision helps ensure this by identifying potential blind spots not directly revealed by the cases.

On the positive side, many of the new sectors are likely to be adequately covered with the external forces already identified in BYTE cases. Because these sectors face similar external forces, they are also likely to benefit from similar policy levers. The new sectors which are adequately covered include hospitality and household activities; these are likely well-covered by the factors considered in the health care case. Less obvious is the real estate case, but our respondents feel that this sector is being shaped by very similar forces to health care. Similar policy responses to personal data, location data, demographic forces, shared and individual risks and responsibilities may well unify both health and real estate.

Likewise we feel that the agriculture sector is adequately covered by the shipping case work. Both sectors involve massive bulk handling and transport, and benefit from global logistics networks and local and global sensor and imagery networks. It should be noted that this standard industrial code also includes mining and forestry as well. It is important to acknowledge that, even if the macro forces shaping the big data future are well-covered by existing case work, there are other contextual factors well worth more detailed surveys in agriculture, forestry and mining.

We believe that many of the forces affecting the new service industry are adequately covered in the smart cities case. Likewise the manufacturing and professional services sectors are adequately covered by the culture, and the oil and gas cases. Downstream operations in the oil and gas sector often entail bulk manufacturing activities. Likewise culture and professional sectors often involve highly creative work. The standard industrial code for professional services also involves scientific and managerial services.

What perhaps has not been adequately covered by the BYTE case work are the sectors of construction, infrastructure, wholesale and retail trade, finance (including insurance) and education. Additional anticipatory governance work, building on the health care case could help extend the work from health care to real estate. As discussed earlier there are definite parallels in the exposure to external forces across these sectors. Anticipatory work in real estate could also help extend policy planning to the construction sectors.

Still, as noted in the BYTE vision, European sectors are too little prepared for a high technology, but open future. This leaves four major sectors of European activity exposed to fundamental environmental uncertainty. The under exposed sectors are finance and insurance, infrastructure, wholesale and retail trade, and education. The roadmapping exercises of work package 6 will help address these potential blind spots for Europe by means of capability planning and research and development planning exercises

7 DISCUSSION AND CONCLUSIONS

This document described the vision statement for the BYTE project. This vision consists of three trends shaping the big data future for Europe, and a set of scenarios. The scenarios reveal that big data externalities potentially impact a wide variety of significant European sectors. Furthermore of the sectors investigated by the BYTE project, many seem ill-prepared for a set of high-technology and open data futures.

Ultimately the BYTE vision entails supporting decision-making to create a Europe which will respond in a more agile, and adaptive manner, to a range of possible external forces. It is not clear which of the futures presented in the vision will be the most satisfactory for Europe. The futures are clearly multi-objective and multi-actor. Different European stakeholders will naturally desire a range different outcomes for Europe as a whole.

Deliverable D5.1 takes up this challenge of describing the outcomes of interest for a variety of different European actors when it comes to formulating European policies. This deliverable also helps inventorize a range of policy levers which may be applied when shaping European outcomes. These levers will be staged and sequenced in the roadmapping activities of workpackage 6.

In concluding this document it is important to acknowledge why big data is a problem for European policy in a single statement. A problem statement acknowledges the emergence of future trends, and describes why the current system is underperforming. The policy system can be underperforming either because it fails to achieve its full and desired potential, or because the system introduces additional undesirable impacts.

Given the concensus statement derived from multiple cases, workshops, and the cross-impact analyses, the following large-scale problems in European big data policies are apparent.

1. Given the rate of technological change in big data, European policy setting may be partially unprepared for the positive and negative impacts resulting from a technological transition towards big data. Should the transition be slower than expected, policy setting should do no harm.
2. Given the political economy of big data operations, European policy setting may be poorly equiped for changes in the hegemony of big data. If a hegemony of a few big external public or private players emerge in big data, Europe needs to exert its influence to hedge or shape the big data future. Alternatively, Europe needs to come to grip with potential futures where a diverse big data ecology is fully established.
3. Given the regime of big data operations, European policy setting needs to be prepared to address both open and public data sources, as well as closed and proprietary protections on data. In particular, many private European sectors are poorly prepared to transition to a potential expansion of the use of open big data.

References

Barber, M., 2009. Questioning scenarios. Journal of Futures Studies 13, 139-146.

Bezold, C., 2009. Jim Dator's alternative futures and the path to IAF's aspirational futures. Journal of Futures Studies 14(2), 123-134.

Bezold, C., 2010. Lessons from using scenarios for strategic foresight. Technological Forecasting and Social Change 77, 1513-1518.

Bishop, P., Hines, A., Collins, T., 2007. The current state of scenario development: An overview of techniques. Foresight 9(1), 5-25.

Borjeson, L., Hojer, M., Dreborg, K-H., Ekvall, T., Finnveden, G., 2006. Scenario types and techniques: towards a user's guide. Futures 38, 723-739.

de Brabandere, L., Iny, A., 2010. Scenarios and creativity: thinking in new boxes. Technological Forecasting and Social Change 77, 1506-1512.

Bradfield, R., Wright, G., Burt, G., Cairns, G. and Van Der Heijden, K., 2005. The origins and evolution of scenario techniques in long range business planning. Futures 37, 795-812.

Burchsted, S., Crews, C., 2003. Systemic scenarios: creating synergy through scenarios and systems thinking. unpublished Master's thesis, University of Houston-Clear Lake, Clear Lake, TX.

Burt, G., van der Heijden, K., 2003. First steps: towards purposeful activities in scenario thinking and future studies, Futures 35, 1011-1026.

Burt, G., 2010. Revisiting and extending our understanding of Pierre Wack's the gentle art re-perceiving. Technological Forecasting and Social Change 77, 1476-1484.

Cappert, C., 1973. The development of a pattern model for social forecasting. Futures 6(4), 367-382.

Chermack, T.J., Lynham, S.A., Ruona, W.E.A., 2001. A review of scenario planning literature, Futures Research Quarterly 17, 7-31.

Chung, J., Linder, J., Liu, I., Seltzer, W. and M. Tse (2016), "Democratic Structures in Cyberspace," Whitepaper 6.805-10, http://groups.csail.mit.edu/mac/classes/6.805/student-papers/fall98-papers/democracy/whitepaper.doc

Coyle, R.G., Crawshay, R., Sutton, L., 1994. Futures assessment by field anomaly relaxation Futures 26(1), 25-43.

Curry, A., 2009. From foresight to insight: using scenarios well. Journal of Futures Studies 13, 119-122.

Curry, A., Schultz, W., 2009. Roads less travelled: Different methods, different futures. Journal of Futures Studies 13(4), 53-60.

Dator, J., 1979. The futures of cultures and cultures of the future. In (eds.) Marsella, T., Ciborowski, T., Tharp, R., Perspectives on Cross Cultural Psychology. Academic Press, New York, NY.

Dator, J., 2002. Advancing futures: Futures studies in higher education. Praeger, Westport, CT.

Dator, J., 2009. Alternative futures at the Manoa school. Journal of Futures Studies 14(2), 1-18.

Douglas, M., 1982. In the active voice. Routledge and Kegan Paul, London, UK.

Eisenhardt, K. M. 1991. Better stories and better constructs: The case for rigor and comparative logic. Academy of Management Review, 16, 620-627.

Eisenhardt, K. M., Graebner, M. E. 2007. Theory building from cases: Opportunities and challenges. Academy of Management Journal, 51(1), 25-32.

Fontela, E., Hingel, A., 1993. Scenarios on economic and social cohesion in Europe, Futures 25, 139-154.

Futures Wheel, Futures Research Methodology Version 3.0, The Millennium Project, Washington, DC 2009.

Godet, M., 2000a. Forefront: how to be rigorous with scenario planning. Foresight 2, 5-9.

Godet, M., 2000b.The art of scenarios and strategic planning: tools and pitfalls, Technological Forecasting and Social Change 65, 3-22.

Grunwald, A., 2011. Energy futures: diversity and the need for assessment. Futures 43, 820-830.

Hiltunen, E., 2009. Scenarios: process and outcome. Journal of Futures Studies 13, 151-152.

Inayatullah, S., 2009. Six pillars: futures thinking for transforming. Foresight 10(1), 4-21.

Joseph, C.F., 2000. Scenario planning. Technological Forecasting and Social Change 65, 115-123.

Kahn, H., Wiener, A.J., 1967. The year 2000: A framework for speculation on the next thirty-three years. The Macmillan, New York.

Lempert, R. J., S. W. Popper and S. C. Bankes (2003). Shaping the Next One Hundred Years : New Methods for Quantitative, Long-term Policy Analysis. Santa Monica, CA, RAND Corporation.
Martelli, A., 2001. Scenario building and scenario planning: state of the art and prospects of evolution. Futures Research Quarterly 17, 57-70.
Martino, J.P., 2003. A review of selected recent advances in technological forecasting, Technological Forecasting and Social Change 70, 719-733.
Pasquale, F. (2015), The Black Box Society: The Secret Algorithms That Control Money and Information, Cambridge, MA: Harvard University Press.
Pillkahn, U., 2008. Using Trends and Scenarios as Tools for Strategy Development. Publicis Corporate Publishing, Erlangen, Germany.
Rhyne, R., 1974. Technological forecasting within alternative whole futures. Technological Forecasting and Social Change 6, 133-l62.
Rhyne, R., 1981. Whole-pattern futures projections using field anomaly relaxation. Technological Forecasting and Social Change 19, 331-360.
Redman, T. C. (2013) "Invest in Proprietary Data for Competitive Advantage," Harvard Business Review, 28 March 2013, accessed online: https://hbr.org/2013/03/invest-in-proprietary-data-for
Schoemaker, P.J.H., 1991. When and how to use scenario planning: A heuristic approach with illustration. Journal of Forecasting 10, 549-564.
Schoemaker, P.J.H., van der Heijden, C.A.J.M., 1992. Integrating scenarios into strategic planning at Royal Dutch/Shell. Strategy & Leadership 20, 41-46.
Schoemaker, P.J.H., 1993. Multiple scenario development: its conceptual and behavioral foundation. Strategic Management Journal 14, 193-213.
Schultz, W., 1993. Scenario building: the Manoa approach. www.infinitefutures.com/tools/sbmanoa.shtml (accessed 19/9/2015)
Siggelkow, N. 2007. Persuasion with case studies. Academy of Management Journal 51(1), 20-24.
Strauss, J.D., Radnor, M., 2004. Roadmapping for dynamic and uncertain environments. Research Technology Management 47, 51-57.
United Nations, 2008. International standard industrial classification of all economic activities (revision 4). Department of Economic and Social Affairs Statistics Division
Van Notten, P., Rotmans, J., van Asselt, M., Rothman, D., 2003. An updated scenario typology. Futures 35, 423-43.
Varum, C.A., Melo, C., 2010. Directions in scenario planning literature. A review of the past decades. Futures 42,355-369.
Wack, P., 1985a. Scenarios: Shooting the rapids. Harvard Business Review 63(5), 139-150.
Wack, P., 1985b. Scenarios: Uncharted waters ahead. Harvard Business Review 63(5), 73-89.
Wildavksy, A., 1987. Choosing preferences by constructing institutions: A cultural theory of preferences formation. American Political Science Review 81(1), 3-22.
Wilkinson, A., 2009. Scenarios practices: in search of theory. Journal of Futures Studies 13, 107-114.
Wood, C.W., Christakis, A.N., 1984. A methodology for conducting futures-oriented workshops. Technological Forecasting and Social Change 26, 281-297.