Requirements Engineering Visualization: A Survey on the State-of-the-Art

John R. Cooper Jr., Seok-Won Lee, Robin A. Gandhi and Orlena Gotel
University of North Carolina at Charlotte, North Carolina, U.S.A.
University of Nebraska at Omaha, Nebraska, U.S.A.
Pace University, New York, U.S.A.
{jrcooper,seoklee}@uncc.edu, rgandhi@unomaha.edu, ogotel@pace.edu

Abstract

Requirements engineering visualization is a rapidly growing field of research; however, the specific characteristics of what makes for effective visualizations during a particular engineering phase have not yet been distinguished. Visualizations, when coupled with traditional practices, augment the ability of resulting requirements artifacts to reach a wide range of stakeholders and provide for a rapid and shared understanding of complex information. This paper represents a survey of the research papers presented during the REV workshops from 2006 to 2008 in order to ascertain how the research trends have evolved over the past few years. By examining approaches to requirements engineering visualization that have been proposed, in retrospect, we hope to show the areas of recent focus, as well as to discover those areas that may hold opportunities for further research with respect to the most commonly understood RE lifecycle phases and activities. In the process, we offer a preliminary classification scheme through which to categorize the various research efforts. Where none existed before, the resulted categorization enables a constructive discussion about the coverage of previous REV contributions from various perspectives, while discovering the gaps, and provides opportunities for further research with the understanding of the trends of applying visualization in requirements engineering research and practice.

1. Introduction

The activities performed during Requirements Engineering (RE) are of the utmost importance to the successful design and implementation of software systems. These early activities in the software development process represent a time when business goals are transformed into technical specifications, when conflicting goals put forth by various stakeholders are negotiated into an agreed upon vision of the software, and when analysis and revision can be performed without the risk of incurring the serious time or budget overruns that can occur as a project reaches more advanced stages of development. Traditional, primarily textual, specification documentation can be cumbersome for many stakeholders involved in the early phases of a project and, consequently, may present obstacles when attempting to come to a shared understanding of the proposed system. Similar challenges exist when reviewing and understanding requirements artifacts for large and complex systems or legacy systems. In addition, certain characteristics of requirements or the problem domain may only become apparent upon visual inspection of the elicited information or metrics.

Basic visualization techniques that are widely accepted in both business and society have long been used in RE. These techniques typically include bar graphs, pie charts, and hierarchical structures. These visualizations are routinely used to augment textual requirements with summarization that aggregate large amounts of information into a single representation for shared understanding and quick absorption by stakeholders. However, requirements engineering visualization is fundamentally challenging due to the typical lack of structure in those artifacts that are created and available during the very early stages of system development. It needs to address the hard problem of transforming informal sources in the problem space into a visual representation that relies on the existence of well-defined metrics that modulate its structural characteristics. As a result, use of visualization is therefore most attractive to overcome the challenges presented by traditional natural language requirements, though in some cases visualization has also been suggested to convey more complex concepts such as the “health” of a set of requirements [21].
Research in the area of visualization pertaining to the activities of RE has received relatively little formal attention until recent years. The First Int’l Workshop on Requirements Engineering Visualization (REV) was held in 2006. This is somewhat surprising considering that the breadth of the activities involved in RE is the most content rich of all the phases of the software development lifecycle. Gotel et al. describe it best in [29]:

*The most data intensive and media-rich aspects of software engineering are clearly those early requirements engineering activities in which stakeholders are determined, problems explored and goals defined, so the period in which informal aspirations converge to an agreed statement of the problem and requirements specification.*

The complexity of systems and the rich socio-technical contexts that significantly challenge present-day RE practices demand the increased use of visualization techniques in order to better capture the rationale for, and specification of, software systems. As more and more visualization techniques become available there is a need to start understanding their suitability for the task at hand as well as identify their ability to complement each other such that multiple perspectives of the problem domain can synergistically influence decision making. While growing attention to this field reveals the perceived potential value in developing visualization techniques and artifacts, to date there is still no clear sense of the specific characteristics that make such artifacts of added value to the process. Therefore, we attempt to reflect upon the fundamental problem of determining which visualization techniques best fit which kind of phases and activities in a typical RE process through a retrospective analysis of the various contributions to date. The ability to gain such perspective was a topic of considerable interest and discussion among participants of REV 08 to advance this rapidly growing field. We believe that REV, currently in its 4th year, is in need of instruments that can be used for describing and somewhat measuring its outcomes, such that its future contributions can be effectively planned and evaluated.

Towards our goal, we have engaged in a survey of the recent activities in the field of requirements engineering visualization and expect the results to provide direction for those areas of visualization research that should be explored further. In particular, we focus on the research papers presented at the REV workshops over the past three years (2006 to 2008) and attempt to discover the trends with respect to the aspects of the RE lifecycle that have been attended to and supported. We expect such trend analysis to clearly distinguish those areas in which the research has been the most focal from those areas that are yet to be explored and so possibly worthy of more investigation.

The rest of this paper is organized as follows: Section 2 describes the research approach used in our survey of REV research contributions. In Section 3, we present the results of our survey and its mapping to the RE lifecycle and visualization types that we classify. We discuss our findings in Section 4. Finally, in Section 5 we summarize our contribution and propose a direction for future efforts in this field.

2. Research Approach

To conduct a survey of the first three years of the REV workshops we undertook four steps:

1. Established a unified perspective of the RE lifecycle in terms of representative phases and activities. This step was based on pre-existing work [30].
2. The primary focus of each paper presented in the past REV workshops was identified and mapped to the most relevant phases and activities in the unified perspective of the RE lifecycle.
3. The types of visualization artifacts most prominently employed within each paper were identified, while attempting to discover correlations pertaining as to how these artifact types applied to the specific RE lifecycle phases and activities (as per Step 2).
4. The trends in the RE lifecycle focus were examined along with use of the different visualization types over the three year period of REV publications.

2.1. Analysis to Map to the RE Lifecycle

To discuss and frame REV contributions within the scope of RE, a general classification scheme is introduced. From the RE-centric visualization perspective, it seems natural and befitting to relate the visualizations to the corresponding progression of RE artifacts (starting with context and groundwork to structured specification, evolution and maintenance) and the activities that guide this progression (starting with elicitation to verification and validation). This categorization also enables us to examine the visualization goals, users (writers and readers of visualizations), achievements, and impact in the
context of requirements engineering purposes, stakeholders, outcomes, and benefits.

Therefore, the first step in our survey process was to baseline a representation of the RE lifecycle process that separates phase and activity dimensions to depict the provenance and flow of artifacts through an intertwined set of RE phases and activities. Many research efforts have characterized the RE process in different ways [32][33][34]. We were able to discover fundamental similarities across these characterizations when mapped to phase and activity dimensions of the RE lifecycle.

RE Phases:

**A: Context and Groundwork**
- Establish the business case.
- Scope the system.
- Mitigate serious risks.
- Establish process, methods, and techniques.
- Assess feasibility.

**B: Requirements Elaboration**
- Prepare initial system model.
- Document high-level organizational needs.
- Gather stakeholder needs and constraints.

**C: Requirements Refinement**
- Original artifacts are refined.
- Interactions among diverse artifacts are identified.
- Conflicts among requirements are resolved.
- Stakeholders agree on a set of requirements for the system.

**D: Requirements Specification, Management and Evolution**
- Precise software specifications are produced from the artifacts.
- Ensure readability and traceability of requirements.
- Document change, or need for change is managed.
- Modifications to accommodate corrections, environmental changes, or new objectives.

In particular, we mapped the characterizations of the RE process provided in [32][33][34] to four distinct categories along the phases and activity dimensions each. These characterizations are elaborated in the above boxes.

The phases and activities described here are visually depicted in Figure 1 (originally derived from [30]). Motivated by the unified software process model [31], Figure 1 shows a unified view of how time and activity dimensions are emphasized within the RE lifecycle. The phases are listed horizontally and indicate the stages of RE over time, advancing from left to right. Similarly, the activities associated with RE are listed from top to bottom along the left side of the figure. The iterations in Figure 1, explicitly
demonstrate that the RE activities are intertwined with each other and hence we continuously iterate through them as we progress along the phases dimension. Finally, for easy reference the activities are assigned labels 1 to 4, from top to bottom, while phases are assigned the letters A to D, from left to right.

In order to map the papers presented during the REV workshops to RE phases and activities, a careful examination was hence undertaken of each research paper by a team of four researchers (of the authors of this paper, two were REV program committee members). Independently, each researcher followed the steps identified in the beginning of Section 2 to identify the appropriate RE phases and activities in which the proposed visualization discussed in the paper could best be applied. Each researcher was presented with a collection of 29 REV papers, a description of the RE phases and activities, and a detailed set of process steps to determine a category into which to map and so assign the paper. Note that 4 phases and 4 activities led to 16 possible categories.

In order to support the mapping process, the RE phases and activities were broken down to emphasize certain characteristics and criteria associated with each, and an elaboration of Figure 1 was provided to each researcher. Each researcher was also provided with a classification of visualization types (see Section 2.2) to assist further in the mapping of papers to visualization types. Each researcher worked independently and the results of the process were later consolidated.

### Table 1. Sample Analysis of [12]

<table>
<thead>
<tr>
<th>Line of Text</th>
<th>Phase</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>…we propose a methodology involving maximum participation of stakeholders</td>
<td>B</td>
<td>1</td>
</tr>
<tr>
<td>where the soft goals are elicited in group sessions iteratively.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In this article we propose an agile visualization technique for eliciting soft</td>
<td>C</td>
<td>1</td>
</tr>
<tr>
<td>goals by doing Agent Based Goal Refinement (ABGR) process repeatedly.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>During the “Develop” sub-activity ABGR process is taken up to elicit soft</td>
<td>B/C</td>
<td>1</td>
</tr>
<tr>
<td>goals… for “Reviewing” by all stakeholders</td>
<td></td>
<td></td>
</tr>
<tr>
<td>During the last sub-activity all stakeholders agree and “Adjust” their own</td>
<td>C</td>
<td>3</td>
</tr>
<tr>
<td>elicited goals…</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Analysis of each paper was typically performed through keyword searches and by careful reading in order to determine how the research goal applied to the sixteen corresponding categories of the RE lifecycle shown in Figure 1. Each paper could be placed in as many as two categories by each researcher in order to account for the inevitable overlap that occurs among the phases and activities of the RE lifecycle. In order to be counted in a category after consolidating all four independent results, a paper had to have been placed in that category independently by at least two of the four researchers. Note that this is a task that would have been simplified greatly if each REV paper had explained explicitly what aspect of RE it set out to support; this was very rarely the case.

As an example of the process, consider the paper: “A Visualization Technique for Agent Based Goal Refinement to Elicit Soft Goals in Goal Oriented Requirements Engineering” by A.M. Sen and S.K. Jain in the proceedings of REV 2007 [12]. In Table 1, a number of pertinent lines from the text of the paper are given alongside the corresponding RE phase or activity that was interpreted to be the most relevant by one researcher. Through this type of analysis, it was determined by one researcher that this paper was most relevant to the phases B and C, and activity 1 (for brevity we refer to them as B1 and C1).

### 2.2. Analysis to Map to Visualization Type

In addition to mapping each REV paper into a category of the RE lifecycle, we also attempted to categorize the primary visualization artifact of each paper. In order to do this, we identified five generic categories of visualizations that provided sufficient coverage of the past REV papers:

- **Tabular visualizations.** Tabular visualizations are made up of a series of intersecting rows and columns that typically hold textual information.
- **Relational visualizations.** Relational visualizations consist of a collection of nodes and connectors that describe or indicate a relationship between components or a system, but do not implicitly describe the inherent order of operation of the system.
- **Sequential visualizations.** Sequential visualizations, on the other hand, convey the order of operation between parts of the system, or of a user and the system. These are akin to traditional flow charts or state-machine diagrams.
- **Hierarchical visualizations.** Hierarchical visualizations imply the decomposition of a system and its parts, as typically used in goal-based modeling approaches.
- **Quantitative/Metaphorical visualizations.** Quantitative visualizations are most commonly seen in the form of bar graphs, pie charts or other figures that convey relative data, but also include more sophisticated techniques that make use of visual metaphors and other visual clues such as color, shape, line thickness, and size to convey meaning at a glance.

### 2.2.1. Analysis to Map to Visualization Type

<table>
<thead>
<tr>
<th>Line of Text</th>
<th>Phase</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>…we propose a methodology involving maximum participation of stakeholders</td>
<td>B</td>
<td>1</td>
</tr>
<tr>
<td>where the soft goals are elicited in group sessions iteratively.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>In this article we propose an agile visualization technique for eliciting soft</td>
<td>C</td>
<td>1</td>
</tr>
<tr>
<td>goals by doing Agent Based Goal Refinement (ABGR) process repeatedly.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>During the “Develop” sub-activity ABGR process is taken up to elicit soft</td>
<td>B/C</td>
<td>1</td>
</tr>
<tr>
<td>goals… for “Reviewing” by all stakeholders</td>
<td></td>
<td></td>
</tr>
<tr>
<td>During the last sub-activity all stakeholders agree and “Adjust” their own</td>
<td>C</td>
<td>3</td>
</tr>
<tr>
<td>elicited goals…</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Of the entire set of papers that were included in this survey, there were a few whose visualization artifacts could not be suitably included within our five classifications. For example, the *i* framework, classified as primarily relational for the purposes of this paper ([2], [11] and [13]), represents a unique blend of many of the elements listed above and may deserve to be mentioned separately in future refinements. Also, the use of visual analogies has been noted as a unique category in [21].

3. Visualization of Results

A total of 29 papers were reviewed from the REV workshops held between 2006 and 2008. While the sample size may appear to be small, three years is a considerable amount of time to study a rapidly advancing requirements visualization field. Papers that did not directly contribute visualization artifacts [3],[29] or apply to the RE process [8],[9],[19] were not included in our survey.

Figure 2 presents the consolidated results of our survey analysis. The matrix outlined in this figure corresponds to the RE phases and activities of Figure 1. A reference to a paper that is placed in a cell of this matrix indicates that there was an agreement in the mapping of the paper to the category between two or more of the researchers. Note that twelve papers have been mapped to two distinct categories in Figure 2, and three papers have been mapped to three distinct categories, in each case indicating agreement of multiple assignments by the researchers.

Below the reference to a paper in Figure 2, the shape identifies the primary type of the visualization artifact that is described in the paper. The use of shapes offers a more detailed view of the distribution while still allowing for the discerning of potential trends. In Figure 3, the mappings are separated by year, each diagram representing the papers presented during a single year’s REV workshop.

4. Findings and Discussion

Our best efforts were made to accurately map the research onto the most appropriate of the sixteen categories in the RE lifecycle; however, not every
paper dealt directly with the creation of visualization artifacts and relevant parts of the software development process (i.e., RE). Instead, some research focused on support for visualization techniques through automation, algorithm refinement, or process improvement. Due to the overlapping nature of the phases and activities within RE, and considering that many processes are reliant on one another in an iterative nature, it was often difficult to precisely identify single categories for classification purposes. We tried to address this problem, to some extent, by allowing papers to be mapped into more than one phase or activity as described in Section 2.1. Likewise, the visualization artifacts employed by or produced from each research paper were not always limited to just one of the five types as identified in Section 2.2. Our efforts sought, instead, to identify the visualization type of the artifact that was the most prominent in the paper or the most important to the overall research effort.

Due to the small sample size of the research in this area, both overall and year-to-year, the current paper does not try to suggest implication for the fitness of the visualization types to a particular RE phase or activity. Equally, a lack of an a priori identification of the primary RE lifecycle focus and their types of visualization artifact of each paper by the authors themselves, may limit the significance of our categorization effort; the categorization represents only the opinions of the researchers who participated and not the original paper authors.

Nevertheless, the findings drawn from undertaking the research to create Figures 2 and 3 suggest that, even in a short amount of time (i.e., three years), there have been innovations in the field of requirements engineering visualization that could prove to be of great worth when used in conjunction with traditional RE techniques and documentation. Due to the lack of time that has been allocated to the evaluation of many of these visualizations, the fundamental problem still remains in determining how effectively a particular visualization technique, or characteristic of a technique, applies to the problem context. However, a preliminary analysis of how the types of the visualization artifacts used in each REV paper apply to the stages of the RE lifecycle reveals many insights in this direction, and are discussed below.

Taking a birdseye view of the categorization result in the unified process perspective, it is interesting to observe that the number of papers in each grid section roughly corresponds to the ebb and flow of the four activity traces over time. A more in-depth look at other interesting trends that emerge from the categorization effort is provided below.

First, we present an analysis of each column in the RE lifecycle (i.e., the RE phases).

- Column A (Context and Groundwork) contains two papers, both of which include quantitative visualizations. This may be the most expected result as it reflects a time in the lifecycle when planning, feasibility, and decision-making are the key activities.
- In Column B (Requirements Elaboration), the coverage of the visualization types is most notable. Of the six distinct papers represented here, two contain sequential visualizations, two contain relational, one hierarchical, and one tabular. As this phase reflects a time in the RE lifecycle when requirements elicitation ramps up, it is understandable that diverse visualization techniques would be employed in order to maximize requirements discovery.
- In Column C (Requirements Refinement), the relational visualization type is clearly the most dominant with seven of the nine distinct papers featuring this classification. Although this is probably expected, we are somewhat surprised that the sequential and hierarchical visualization types are not used more during this RE phase as it characterizes a transition from requirements elicitation to specification.
- Once again in Column D (Requirements Specification, Management and Evolution), we see relational visualizations occurring the most frequently, reflected by half of the papers here. Although we see more sequential, hierarchical, and quantitative visualizations in Column D, we expected even more attention as the requirements specification phase incorporates the increased need to define system structure and to structure the requirements so as to manage change and evolution.

Finally, when examining the predominant visualization type across columns, and specifically the relationship between columns B, C, and D, we see the prevalence of relational and hierarchical visualizations. This possibly reflects the fact that requirements structure is being established in these RE phases. Conversely, it appears that the number of visualizations being used to depict sequencing and ordering information is diminishing. Perhaps this could offer some insight as to why traceability back to the provenance of requirements is often so difficult as the RE lifecycle progresses?
A similar analysis was undertaken as to how the REV papers were mapped to rows in the RE lifecycle (i.e., the RE activities). Overall, it is evident that visualization techniques for the main activities of RE have received much coverage.

- Row 1 (Elicitation, Understanding, and Structuring) contains a diverse array of visualizations, and they reflect the need to structure and quantify requirements in the formative RE activities.
- Row 2 (Modeling and Analysis) contains the majority of the visualization efforts, but these are predominantly of the relational type.
- An overall look, combining Rows 2 (Modeling and Analysis) and 3 (Communication and Negotiation), the large number of visualization efforts is representative of what is almost certainly a typical bias in the view of how visualizations can best be used during RE; specifically, how visualization techniques are probably best applied to support the efforts of requirements modeling, analysis, communication, and negotiation.
- In stark contrast is the fact that, considering Rows 1 (Elicitation, Understanding, and Structuring) and 4 (Verification and Validation) together, we see only three papers [10, 12, 16] that are unique to either of these rows, i.e. these papers are not also relevant in some other phase and activity. Surely, this represents an opportunity for future research in expanding the role of visualization techniques exclusively in these areas, especially to support

Figure 3: Analysis of Visualization Contributions in the REV Workshops as Differentiated by Year
the activities of requirements verification and validation in the early phases of RE?

In the longitudinal analysis of REV papers from 2006 to 2008 (Figure 3), the problem of the small sample size is compounded even further and the reader should bear in mind that trends are even more difficult to assign significant weight to. The papers mapped from 2006 yield the least diversity in visualization type, having two sequential, two quantitative, and three relational visualization types. The mappings are also spread over all four of the columns and three of the four rows.

In 2007, tabular and hierarchical visualizations were only evident in two of the REV papers, however, there were no sequential types and only one quantitative type. Relational visualization types were once again the most common and were identified in four of the mappings. The biggest change from 2006 to 2007 appears to occur in the way that the research efforts fall into the phase associated with requirements specification, with a particular focus on verification and validation activities.

Finally, a look at 2008 shows that the focus continued to move in the direction of the later phases of the RE lifecycle, with five of the seven papers being mapped onto either rows 2 or 3 and onto columns C or D. As with the earlier analysis of rows and columns, it is still somewhat surprising that there is no evidence of increased focus in applying quantitative or hierarchical visualizations to these later RE phases and their associated activities.

5. Conclusions

Requirements engineering represents a time in the software development lifecycle in which a great amount of information is gathered, refined, communicated, negotiated and tracked from a diversity of different sources, both formal and informal, and among a variety of stakeholders. There is little doubt that, within this rich environment of information, there is the need for visualization techniques that can serve to enhance communication and understanding relative to the essential properties from which software systems will be developed. Effective visualization artifacts can potentially reduce misconceptions and gaps in understanding by presenting many aspects of the RE process, from the health of a set of requirements, to the conflicts between requirements, to requirements traceability, and so on.

Our research has sought to understand the use of visualization techniques in RE during its recent history. Through a study of this nature, we learn more about the characteristics of the various visualization artifacts in relation to how and where they are used to support RE phases and activities. This work is intended as an initial effort in building an evaluation framework regarding the use of visualization in RE. While it is a first attempt at depicting coverage, more research needs to be conducted to explore effectiveness.

In terms of coverage, much opportunity exists to develop visualizations to assist with the pre-requirements phases of understanding the context for RE and for undertaking the groundwork necessary for any RE process. In terms of RE activities, there is much need to develop visualizations to support verification and validation tasks. The dominant types of visualization used for RE can be seen to be relational to date. It is suggested that quantitative/metaphorical visualization types could play a stronger role for those RE phases and activities in which options need to be presented and decisions made. Research will hence lie in developing suitable approaches to extract relative data. More critically, there is certainly a need to better understand the problems and tasks within RE that we are trying to support with visualization and the affordances of the various visualization types in our portfolio. This is an important area in need of more in-depth research exposition.

At this point in the maturation of requirements engineering visualization as a field, there is a need for data to be collected pertaining to the evaluation of recently proposed techniques. Through an analysis of how effective many of the proposed techniques have been over a period of time it will prove easier to determine the specific characteristics inherent to a given technique that makes it more likely to be effective for a particular RE phase or activity. In fact, most of the REV papers concluded that further application of the presented visualization technique would be necessary in order to determine its effectiveness. Through continued innovation in the field, evaluation of current approaches and increased acceptance of non-traditional methods in software engineering, the hope is that these successes propagate throughout the software development lifecycle leading to improved quality in the design and implementation of software systems.

Our future work in this area will continue to address the fundamental question of how certain visualization techniques can be most effectively applied across the RE lifecycle. The next steps will include using the classification to conduct focused compare and contrast exercises among visualization approaches that belong to the same square in Figure 2. In order to facilitate
such effort, a future edition of the REV might propose a framed requirements problem and ask for submissions that evaluate a visualization technique on this problem to allow for through evaluation and comparison. Another suggestion would be to request future REV submissions to explicitly categorize the work based on the phase/activity matrix provided in this paper.

Other aspects of our ongoing work include further refinements to the process of paper classification according to RE phase, RE activity, and visualization type. Finally, through our efforts we hope to offer guidance in the selection of visualization artifacts according to the specific problem context of each RE phase and activity.

6. Acknowledgements
We thank the anonymous reviewers whose valuable comments helped us improve this paper.

7. References


