Scientific Information Management in Collaborative Experimentation Environments
Kaletas, E.C.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
# Contents

List of Abbreviations v

## 1 Introduction

1.1 E-Science Paradigm ........................................ 2
   1.1.1 Definition and Characterization of E-Science Domains and Applications ............... 3
   1.1.2 Target Problem Domains and Applications – Motivation .................................. 5
1.2 Virtual Laboratory Solution ........................................ 9
   1.2.1 Support Environments for Scientific Experimentations .................................. 10
   1.2.2 Virtual Laboratory ........................................ 11
1.3 Research Objectives ........................................ 14
1.4 Research Approach ........................................ 16
1.5 Base Projects ........................................ 19
   1.5.1 The VLAM-G Project ....................................... 19
   1.5.2 Other Related Projects ................................... 22
1.6 Structure of the Thesis ........................................ 23

## 2 Analysis of Requirements for Virtual Laboratories 25

2.1 Approach for Requirements Analysis .................................. 26
2.2 Virtual Laboratory Characterization .................................. 26
2.3 Use Case Analysis ........................................ 33
   2.3.1 Scientists Use Cases ....................................... 34
   2.3.2 Domain Experts Use Cases ................................... 38
   2.3.3 Tool Developers Use Cases ................................... 39
   2.3.4 Administrators Use Cases ................................... 40
   2.3.5 ICT Developers Use Cases ................................... 41
2.4 Classification of Requirements ....................................... 42
2.5 Analysis of User Requirements ....................................... 44
   2.5.1 Scientists Requirements ....................................... 44
   2.5.2 Domain Experts Requirements ................................... 47
   2.5.3 Tool Developers Requirements ................................... 48
   2.5.4 Administrators Requirements ................................... 48
2.6 Analysis of Base ICT Infrastructure Requirements ......................... 49
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.6.1</td>
<td>General VL Requirements</td>
<td>49</td>
</tr>
<tr>
<td>2.6.2</td>
<td>Information Management Requirements</td>
<td>56</td>
</tr>
<tr>
<td>2.7</td>
<td>Conclusions</td>
<td>59</td>
</tr>
<tr>
<td>3</td>
<td>A Framework for Study and Evaluation of Related Work</td>
<td>63</td>
</tr>
<tr>
<td>3.1</td>
<td>Related Enabling Technologies, Standards and Paradigms</td>
<td>65</td>
</tr>
<tr>
<td>3.1.1</td>
<td>Information Models and Standards</td>
<td>66</td>
</tr>
<tr>
<td>3.1.2</td>
<td>Distributed/Federated Information Management</td>
<td>67</td>
</tr>
<tr>
<td>3.1.3</td>
<td>Resource Management Technologies</td>
<td>71</td>
</tr>
<tr>
<td>3.1.4</td>
<td>Other Related Technologies, Paradigms and Tools</td>
<td>75</td>
</tr>
<tr>
<td>3.2</td>
<td>Evaluation Criteria</td>
<td>76</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Identified Requirements Related Criteria</td>
<td>77</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Provided Functionality Related Criteria</td>
<td>79</td>
</tr>
<tr>
<td>3.2.3</td>
<td>Applied Enabling Technologies, Standards and Paradigms Related Criteria</td>
<td>81</td>
</tr>
<tr>
<td>3.2.4</td>
<td>Employed Implementation Technologies Related Criteria</td>
<td>81</td>
</tr>
<tr>
<td>3.3</td>
<td>Classification of Related Work</td>
<td>82</td>
</tr>
<tr>
<td>3.4</td>
<td>Overview of the Studied Systems</td>
<td>84</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Context-Oriented Systems</td>
<td>84</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Computation-Oriented Systems</td>
<td>89</td>
</tr>
<tr>
<td>3.4.3</td>
<td>Instrumentation-Oriented Systems</td>
<td>95</td>
</tr>
<tr>
<td>3.4.4</td>
<td>Metadata-Oriented Systems</td>
<td>97</td>
</tr>
<tr>
<td>3.4.5</td>
<td>Other Systems</td>
<td>101</td>
</tr>
<tr>
<td>3.5</td>
<td>Evaluation of the Studied Systems</td>
<td>101</td>
</tr>
<tr>
<td>3.6</td>
<td>Evaluation of the Proposed VLAM-G/VIMCO Approach</td>
<td>106</td>
</tr>
<tr>
<td>3.7</td>
<td>Conclusions</td>
<td>106</td>
</tr>
<tr>
<td>4</td>
<td>Information Management Framework for CEE</td>
<td>109</td>
</tr>
<tr>
<td>4.1</td>
<td>Modelling Scientific Experiments</td>
<td>111</td>
</tr>
<tr>
<td>4.1.1</td>
<td>Approach Used for Modelling Scientific Experiments</td>
<td>111</td>
</tr>
<tr>
<td>4.1.2</td>
<td>The Proposed Experiment Model</td>
<td>115</td>
</tr>
<tr>
<td>4.2</td>
<td>User Environment of CEE</td>
<td>122</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Facilities Provided by the User Environment</td>
<td>123</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Graphical Presentation of Experiments in the User Environment</td>
<td>124</td>
</tr>
<tr>
<td>4.3</td>
<td>Data Modelling for Information Related to Scientific Experiments</td>
<td>125</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Approach Used for Modelling Experiment-Related Information</td>
<td>125</td>
</tr>
<tr>
<td>4.3.2</td>
<td>Data Modelling for Information about Scientific Experiments</td>
<td>128</td>
</tr>
<tr>
<td>4.3.3</td>
<td>Data Modelling for Information about Users</td>
<td>134</td>
</tr>
<tr>
<td>4.3.4</td>
<td>Data Modelling for CEE-Related Information</td>
<td>137</td>
</tr>
<tr>
<td>4.4</td>
<td>Functionality Modelling for Managing Information Related to Scientific Experiments</td>
<td>138</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Functionality for Managing Information about Scientific Experiments</td>
<td>138</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Policies for Managing Experiment Contexts</td>
<td>145</td>
</tr>
<tr>
<td>4.4.3</td>
<td>Functionality for Managing Scientific Data/Information</td>
<td>153</td>
</tr>
</tbody>
</table>
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5 Functionality Modelling for Collaboration, Administration and Security</td>
<td>154</td>
</tr>
<tr>
<td>4.5.1 Collaboration Functionality</td>
<td>155</td>
</tr>
<tr>
<td>4.5.2 Administration Functionality</td>
<td>156</td>
</tr>
<tr>
<td>4.5.3 Security Functionality</td>
<td>157</td>
</tr>
<tr>
<td>4.6 Methodology for Integrating New Domains and Applications in CEE</td>
<td>160</td>
</tr>
<tr>
<td>4.6.1 Integration of a New Domain in CEE</td>
<td>160</td>
</tr>
<tr>
<td>4.6.2 Integration of a New Application in CEE</td>
<td>161</td>
</tr>
<tr>
<td>4.7 Conclusions</td>
<td>162</td>
</tr>
<tr>
<td>5 Management of Information in the VLAM-G Experimentation Environment</td>
<td>163</td>
</tr>
<tr>
<td>5.1 VLAM-G Experimentation Environment – Revisited</td>
<td>164</td>
</tr>
<tr>
<td>5.1.1 Experiment Model of the VLAM-G</td>
<td>164</td>
</tr>
<tr>
<td>5.1.2 VLAM-G User Environment: Front-End</td>
<td>164</td>
</tr>
<tr>
<td>5.1.3 Experimentation in the VLAM-G</td>
<td>167</td>
</tr>
<tr>
<td>5.2 VIMCO: Virtual Laboratory Information Management for Cooperation</td>
<td>169</td>
</tr>
<tr>
<td>5.3 The VIMCO Architecture</td>
<td>170</td>
</tr>
<tr>
<td>5.3.1 VIMCO Communication Servers</td>
<td>171</td>
</tr>
<tr>
<td>5.3.2 VIMCO Core Functionality Server</td>
<td>172</td>
</tr>
<tr>
<td>5.3.3 VIMCO DB Servers</td>
<td>177</td>
</tr>
<tr>
<td>5.3.4 Implementation of VIMCO</td>
<td>177</td>
</tr>
<tr>
<td>5.4 VIMCO Databases</td>
<td>178</td>
</tr>
<tr>
<td>5.5 Functionality/Services Provided by VIMCO</td>
<td>183</td>
</tr>
<tr>
<td>5.5.1 Overview of the VIMCO API</td>
<td>184</td>
</tr>
<tr>
<td>5.5.2 Services for Accessing VIMCO</td>
<td>185</td>
</tr>
<tr>
<td>5.5.3 Services for Session, User and Access Rights Management</td>
<td>188</td>
</tr>
<tr>
<td>5.5.4 Services for Managing Experiment-Related Information</td>
<td>192</td>
</tr>
<tr>
<td>5.6 Conclusions</td>
<td>196</td>
</tr>
<tr>
<td>6 The DNA Microarray Application Case</td>
<td>197</td>
</tr>
<tr>
<td>6.1 DNA Microarray Technology and Microarray Experiments</td>
<td>198</td>
</tr>
<tr>
<td>6.2 A Motivating Microarray Experiment Scenario</td>
<td>199</td>
</tr>
<tr>
<td>6.3 The DNA Microarray Application</td>
<td>201</td>
</tr>
<tr>
<td>6.3.1 The Expressive System</td>
<td>201</td>
</tr>
<tr>
<td>6.3.2 Developing and Integrating the Expressive System</td>
<td>201</td>
</tr>
<tr>
<td>6.3.3 Data Loading</td>
<td>214</td>
</tr>
<tr>
<td>6.3.4 Representing Different Types of Experiments</td>
<td>214</td>
</tr>
<tr>
<td>6.4 Using the Expressive System</td>
<td>215</td>
</tr>
<tr>
<td>6.5 Conclusions</td>
<td>216</td>
</tr>
<tr>
<td>7 Conclusions</td>
<td>217</td>
</tr>
<tr>
<td>7.1 Overview of the Proposed Approach</td>
<td>217</td>
</tr>
<tr>
<td>7.2 Summary of Thesis Achievements</td>
<td>219</td>
</tr>
<tr>
<td>7.3 Evaluation and Comparison of the Proposed VLAM-G/VIMCO Approach</td>
<td>221</td>
</tr>
</tbody>
</table>