

# Contents

<b>1</b>	<b>ESM Data Archives in Times of the Grid . . . . .</b>	<b>1</b>
	Reference . . . . .	3
<b>2</b>	<b>Distributed Archives, Databases and Data Portals: The Scene . . . .</b>	<b>5</b>
	References . . . . .	12
<b>3</b>	<b>Harvesting of Metadata with Open Access Tools . . . . .</b>	<b>13</b>
3.1	Background and Motivation . . . . .	14
3.2	The Metadata Portal Framework “panFMP” . . . . .	14
3.3	Adoption of the Framework in C3Grid . . . . .	15
3.4	Conclusion . . . . .	19
	References . . . . .	20
<b>4</b>	<b>Data Discovery: Identifying, Searching and Finding Data . . . . .</b>	<b>21</b>
4.1	Overview and Definition of Terminology . . . . .	21
4.2	ESM Data Discovery Challenges . . . . .	22
4.2.1	Data Product Granularity and Aggregation . . . . .	22
4.2.2	Data Provenance and Scientific Context . . . . .	23
4.2.3	Data Lifetime and Quality . . . . .	23
4.2.4	Data Volume and Data Distribution . . . . .	23
4.2.5	Interdisciplinarity and Heterogeneity of Data Discovery Use Cases . . . . .	24
4.2.6	Search Scalability and Flexibility . . . . .	24
4.3	Generic Data Discovery Architecture . . . . .	24
4.4	Implementation Standards . . . . .	26
4.4.1	Metadata Standards and Initiative: . . . . .	26
4.4.2	Metadata Exchange and Discovery Protocol Standard . . .	27
4.4.3	Standardization of Registries . . . . .	28
4.5	Data Discovery Architectures in the Times of Grid . . . . .	28
4.5.1	NERC Data Grid . . . . .	28
4.5.2	C3Grid . . . . .	29
4.5.3	Earth System Grid (ESG). . . . .	30
4.5.4	The European Distributed ESM Data Archive . . . . .	31

4.6	Summary . . . . .	31
	References . . . . .	32
<b>5</b>	<b>User Driven Data Access Mechanisms.</b> . . . . .	<b>33</b>
5.1	Existing Data Access Solutions in ESM . . . . .	34
5.2	Technical Solutions for Data Access . . . . .	37
5.3	Aspects of Local Implementations . . . . .	41
5.4	Future Developments . . . . .	43
5.5	Conclusions . . . . .	45
	References . . . . .	45
<b>6</b>	<b>Collaborative Climate Community Data and Processing Grid—C3Grid: Workflows for Data Selection, Pre- and Post-Processing in a Distributed Environment</b> . . . . .	<b>49</b>
6.1	General Remarks . . . . .	49
6.2	Architecture and Middleware Components . . . . .	51
6.2.1	C3Grid Portal . . . . .	52
6.2.2	Data Information Service and Metadata . . . . .	53
6.2.3	Collaborative Workspace and Data Management Service . . . . .	53
6.2.4	Workflow Scheduling Service . . . . .	55
6.2.5	Provider Layer: Data Archive and Compute Resources . . . . .	56
6.3	Workflows . . . . .	57
6.4	Status and Perspectives . . . . .	59
	References . . . . .	60
<b>7</b>	<b>Earth System Grid Federation: Federated and Integrated Climate Data from Multiple Sources.</b> . . . . .	<b>61</b>
7.1	Background . . . . .	61
7.2	Science Drivers . . . . .	65
7.2.1	Typical ESGF Facility: PCMDI LLNL . . . . .	67
7.2.2	Enhancing the Process of Science . . . . .	67
7.3	Data Quality and Publishing in an International Setting . . . . .	71
7.4	Data Intensive Climate Science . . . . .	73
7.5	Data, Workflow, Middleware Tools, and Services . . . . .	74
7.6	Outstanding Issues . . . . .	76
7.7	Summary . . . . .	77
<b>8</b>	<b>Future Perspectives</b> . . . . .	<b>79</b>
	<b>Glossary</b> . . . . .	<b>83</b>
	<b>Index</b> . . . . .	<b>89</b>