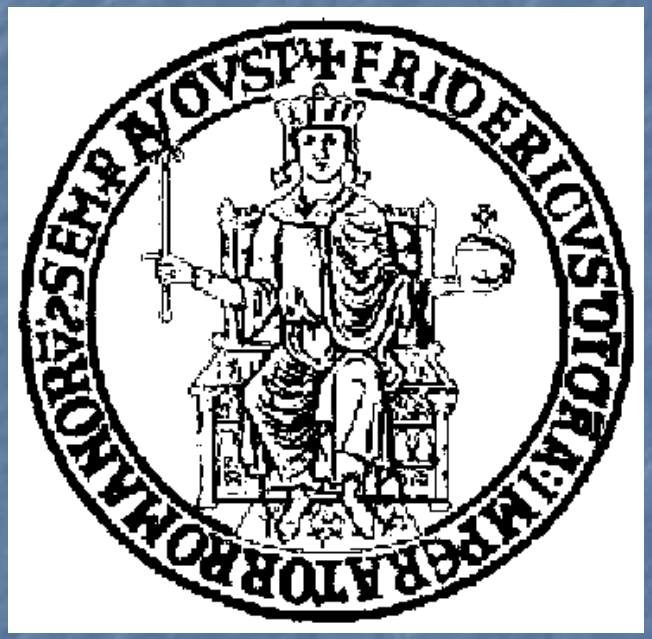
Reverse Engineering Techniques: from Web Applications to Rich Internet Applications



Porfirio Tramontana

Domenico Amalfitano

Anna Rita Fasolino

Dipartimento di Ingegneria Elettrica e Tecnologie dell'Informazione

University of Naples Federico II, Italy

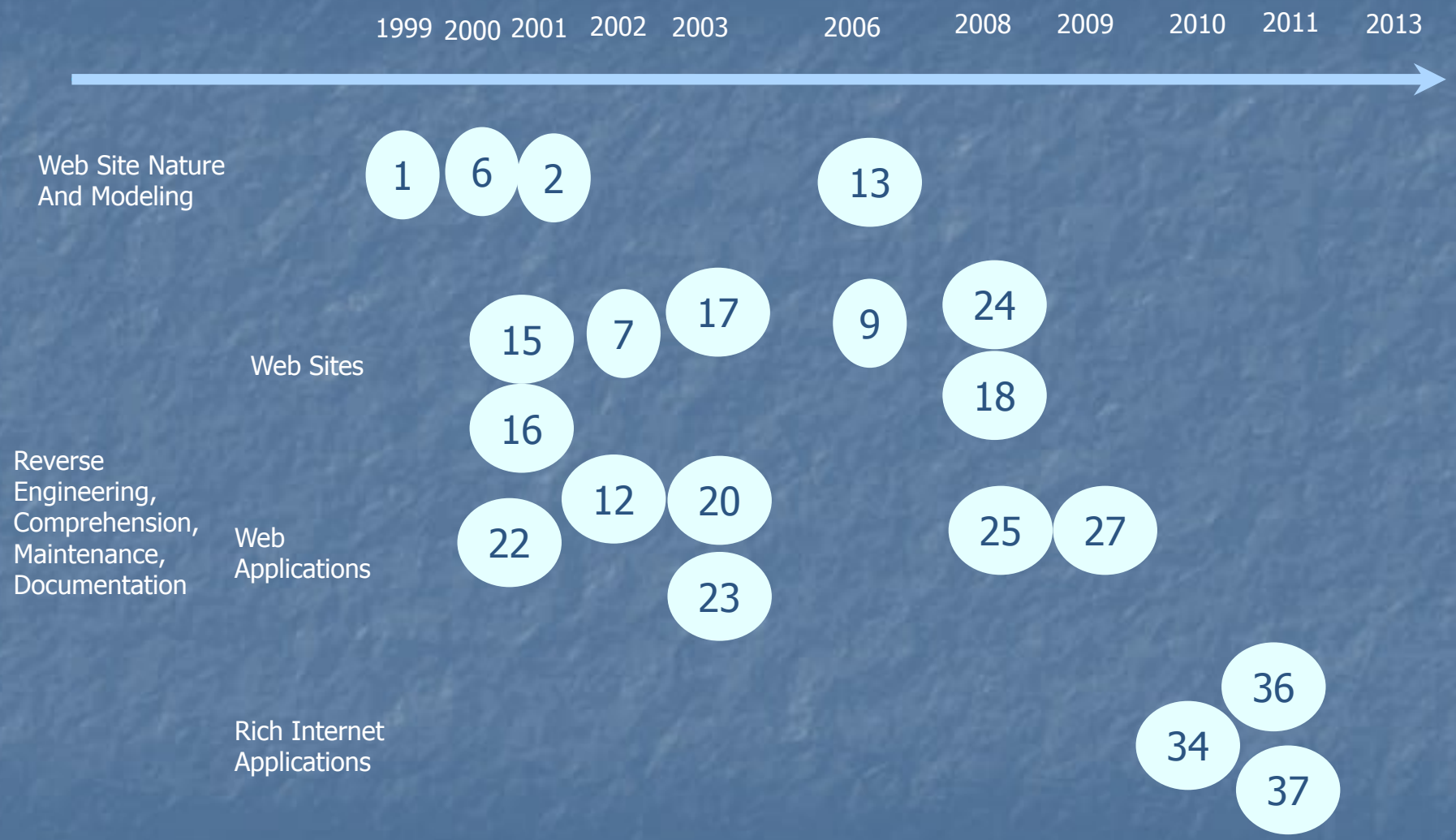
WSE & Reverse Engineering

- Reverse Engineering has been one of the most discussed topics in WSE events
  - Common topics of WSE:
    - Modeling
    - Reverse Engineering, reengineering, refactoring
    - Migration from legacy systems, to Web Applications, to Rich Internet Applications
    - Test case Generation
    - Documentation Generation

WSE & Reverse Engineering

- Web Systems Modeling, 3 papers in 1999-2001 (and a paper in 2006)
- Reverse Engineering, Comprehension, Maintenance, Documentation of:
  - Web Sites, 7 papers in 2000-2008
  - Web Applications, 6 papers in 2000-2009
  - Rich Internet Applications, 3 papers in 2010-2011

WSE Timeline



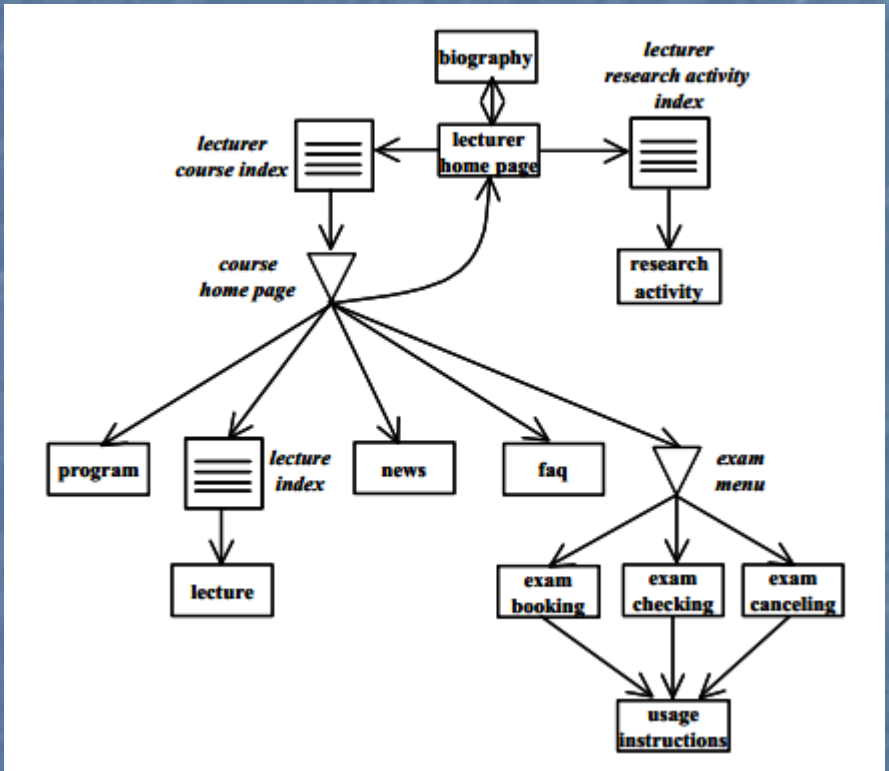
Timeline 1: Web Site Nature and Modeling

- In the first WSE editions, there was discussions about the nature of the Web Sites and the models needed for their design
  - 1999: G. Antoniol, G. Canfora, A. Cimitile, and A. De Lucia, "WEB Sites: Files, Programs or databases?,"
  - 2001: H. M. Kienle and H. A. Muller, "Leveraging program analysis for Web site reverse engineering
- The first models for Web Site design were adaptations of data models (RMM, WebML)
  - 2000: G. Antoniol, G. Canfora, G. Casazza, and A. De Lucia, "Web Site Reengineering Using RMM"
- The evolution of Web Sites to Web Applications caused a corresponding evolution of models towards UML based ones
  - 2006: F. Ricca, M. Di Penta, M. Torchiano, P. Tonella, and M. Ceccato, "An empirical study on the usefulness of Conallen's stereotypes inWeb application comprehension,"

Timeline 1: Web Site Nature and Modeling

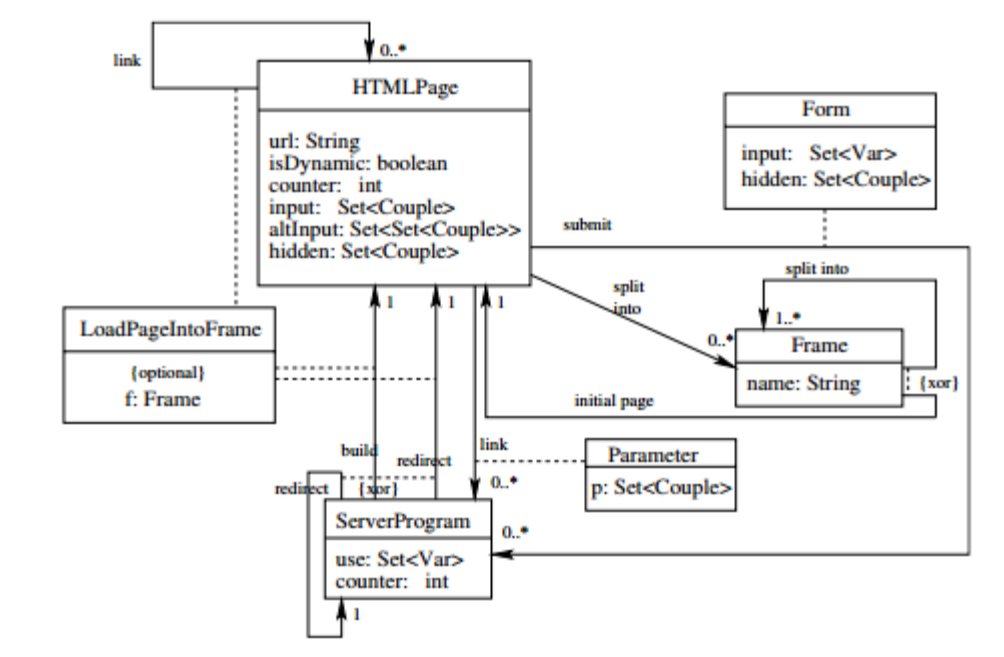
- Reverse Engineering suitable models describe Web Applications at a higher level of details
  - 2002: P. Tonella and F. Ricca, "Dynamic model extraction and statistical analysis of Web applications,"
  - 2002: G. A. Di Lucca, A. R. Fasolino, and P. Tramontana, "Towards a better comprehensibility of web applications: lessons learned from reverse engineering experiments,"
- Models supporting Web 2.0 applications extended the ones suitable for Web applications
  - 2006: R. Djemaa, I. Amous, and A. Hamadou, "WA-UML: Towards a UML extension for modelling Adaptive Web Applications

2000: RMM Model



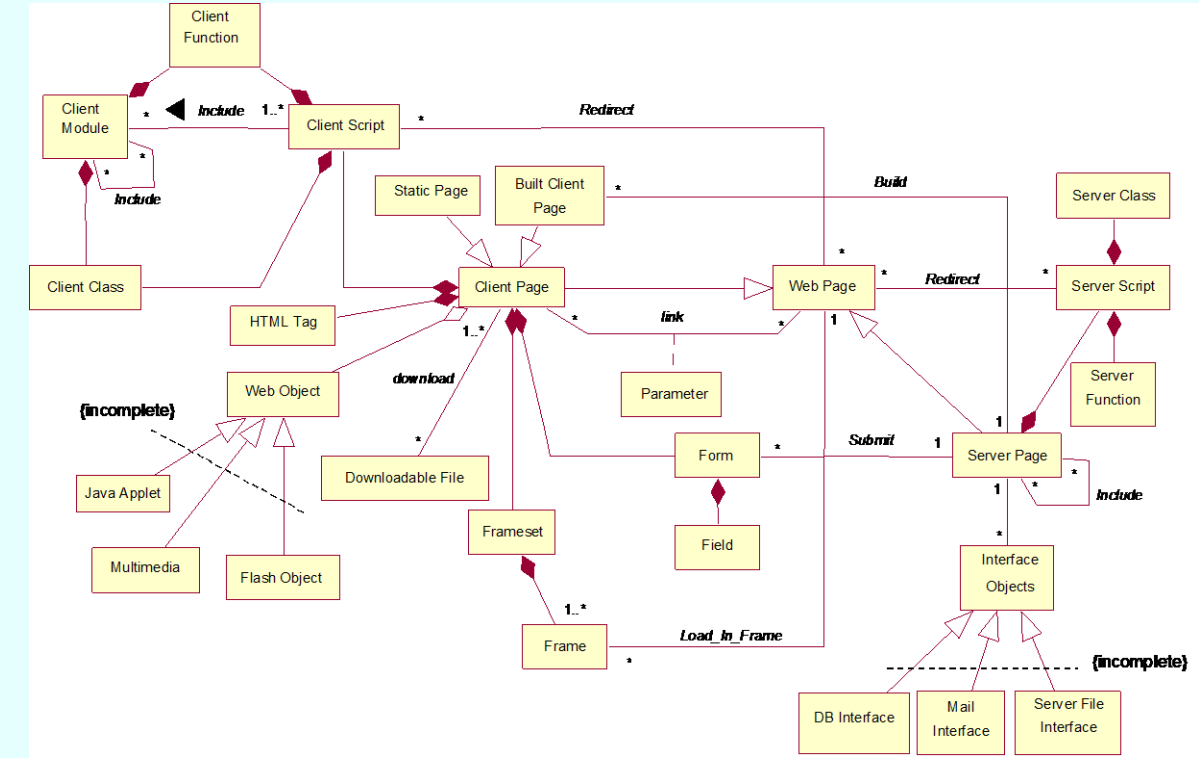
G. Antoniol, G. Canfora, G. Casazza, and A. De Lucia, "Web Site Reengineering Using RMM,"

2002: Dynamic Model



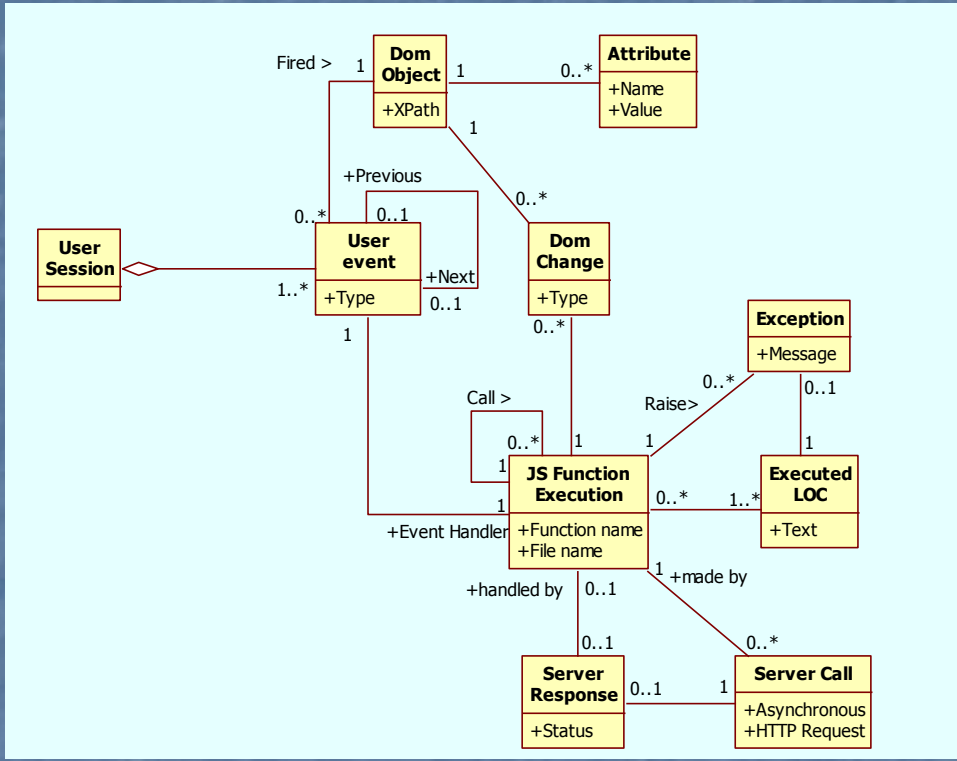
P. Tonella and F. Ricca, "Dynamic model extraction and statistical analysis of Web applications,"

2002: Web Application Model



G. A. Di Lucca, A. R. Fasolino, F. Pace, P. Tramontana, and U. De Carlini, "WARE: a tool for the reverse engineering of Web applications,"

2010: RIA Dynamic Model



D. Amalfitano, A. R. Fasolino, A. Polcaro, and P. Tramontana, "Comprehending Ajax Web Applications by the DynaRIA Tool,"

Timeline 2: Reverse Engineering of Web Sites

- Reverse Engineering for migrating …
  - … from HTML to XML
    - Kirda et al., 2001;
  - … from Web Sites to Web Applications …
    - By abstracting a data model, Estievenart et al, 2003
- Reverse Engineering for reuse …
  - … of clones, Aversano et al., 2001
- Reverse Engineering for reengineering …
  - … based on dynamic analysis and statistic data, Tonella and Ricca, 2002 and 2008
  - … to improve the navigability, Scanniello et al., 2008

Timeline 3: Reverse Engineering of Web Applications

- Part of the 'source' code is generated at run-time
  - Static analysis is not able to recover anything
  - Dynamic analysis is not able to cover anything
- Business logic, GUI and data management are often interleaved

- Di Lucca et al., 2001, 2002, 2003, statically analyzed the source code, abstracted detail level diagrams, reconstructed modular architecture and and abstracted business level UML diagrams
- Hassan and Holt, 2001 extracted architectural diagrams from a combination of static and dynamic analysis
- Ricca et al., 2002, extracted Conallen's diagram from dynamic analysis information
  - Bernardi et al., 2008, et Alalfi et al. 2009, focused on more specific reverse engineering tasks

Timeline 3: References

- 2001: A. E. Hassan and R. C. Holt, "Towards a better understanding of Web applications,"
- 2002: G. A. Di Lucca, A. R. Fasolino, and P. Tramontana, "Towards a better comprehensibility of web applications: lessons learned from reverse engineering experiments,"
- 2003: G. A. Di Lucca, A. R. Fasolino, P. Tramontana, and U. De Carlini, "Abstracting business level UML diagrams from Web applications,"
- 2003: P. Tonella, F. Ricca, E. Pianta, and C. Girardi, "Evaluation methods for Web application clustering,"
- 2006: F. Ricca, M. Di Penta, M. Torchiano, P. Tonella, and M. Ceccato, "An empirical study on the usefulness of Conallen's stereotypes inWeb application comprehension,"
- 2008: M. L. Bernardi, G. A. Di Lucca, and D. Distante, "Reverse engineering of Web Applications to abstract user-centered conceptual models,"
- 2009: M. H. Alalfi, J. R. Cordy, and T. R. Dean, "WAFA: Fine-grained dynamic analysis of web applications,"

Timeline 4: Reverse Engineering of Rich Internet Applications

- RIAs introduced further levels of dynamicity into Web Applications and increasing the difficulties of tasks such as architecture reconstruction and crawling.
  - Asynchronous calls
  - Client side code run-time self-modification
- Pure dynamic analysis approaches have been proposed ...
  - … for test case generation, by Amalfitano et al., 2010
  - … for redocumentation, Amalfitano et al., 2011
  - … for comprehension, McIntosh et al., 2011

Future Perspectives

- Reverse Engineering of Web Applications loses interest because:
  - Web Applications are not more realized from the scratch but their coding is heavily supported by visual tools for code generation and frameworks libraries
    - With Wordpress, CMS, ..., Web applications are essentially 'configured' instead of develop from scratch
  - Web applications are not more so different from other typologies of applications
    - Development models, paradigms and patterns can be the same of traditional applications
      - E.g. a Web Application can represent only a possible user interface for a remote user

Future Perspectives

- But …
  - Many of the static and dynamic analysis techniques initially proposed for Web applications have recently proven their usefulness in the context of mobile applications.
    - E.g.: Android applications are quite similar to RIAs:
      - They are both based on event-based GUIs
      - They are both based on client-server synchronous and asynchronous interactions.