Does Software Engineering Research have Impact on Software Engineering Practice?

A Brief Introduction to the Impact Project

Alexander L. Wolf

Imperial College London

Assumptions

There *is* a software engineering practice

- engineers apply techniques, use tools, and follow processes

There *is* a body of work that can be identified as software engineering research

- both academic and industrial

There *is* skepticism about the impact of research on practice

- *witness:* the relatively low level of academic and industrial research funding *vis-à-vis* the increasingly critical role played by the practice

From where does technology come?

Facile answers are misleading

It comes from…

- Sun, Microsoft, IBM, Google, “the Web”, …

Yes, but!

- from where did they get it? and how?

It comes from…

- Dr. X, who published a seminal research paper and produced a popular prototype

Yes, but!

- we didn’t read the paper
- it was only a prototype, not engineered/licensed for industrial use

It comes from…

- being “in the air”, everybody knows it

Yes, but!

- how did it get there? who nurtured it?

Ideas have many parents…

Researchers

Scientific and technical communities

Technology transfer agents

Students with new degrees

New hires with different perspectives

Early adopters

Commercializers

… and their contributions differ

Initial conceptualization of idea

Evangelism

Prototype demonstration

Public promulgation

Nurturing by community activities

Education, training, and indoctrination

Product commercialization

Why should we care?

Facile answers are misleading

Some technology is not very good

- why are we stuck with it?
- why is it not better?

Some technology seems useful

- how can we get more of it?
- how can we speed its appearance?
- are there approaches that need to be strengthened/nurtured, despite their slow adoption?

How do we evaluate the contributions?

Facile answers are misleading

Ideally

- qualitatively
- quantitatively

But a challenging task

- different parties have different motivations
- apportioning contributions is difficult
- long timescales attenuate measurements and memories

CiP: Pressure to measure in the UK

UK Government is seeking accountability

- demonstrating economic impact of tax-payer investment in basic research and improving exploitation of research outputs

2006 DTI report on "Increasing the Economic Impact of the Research Councils"[†]

- provides several recommendations on how to consider economic impact in funding decisions
- example: an individual competent in the economic impact of research should be accommodated on each review panel

[†] http://www.dti.gov.uk/files/file32802.pdf

But measuring is easier said than done

2006 EPSRC report on "International Review of ICT Research in the UK"[†]

- panel noted the difficulty in conducting macro-economic analysis of ICT commercial impact

2007 Russell Group response to DTI report[‡]

- "There is no evidence to date of any rigorous way of measuring economic impact other than in the very broadest of terms and outputs."

[†] http://epsrc.ac.uk/ResearchFunding/Programmes/ICT/ReviewsAndConsultations/InternationalReview/

[‡] http://www.russellgroup.ac.uk/news/2007/rcuk-consultation-on-the-efficiency-and-effectiveness-of-peer-review.html

Goals of the Impact Project

Scholarly, objective, case-based evaluation

Deliverables

- peer-reviewed papers
- presentation materials and outreach activities
- expertise

Community building

Prospective for future research investment

Lessons learned for "successful" research

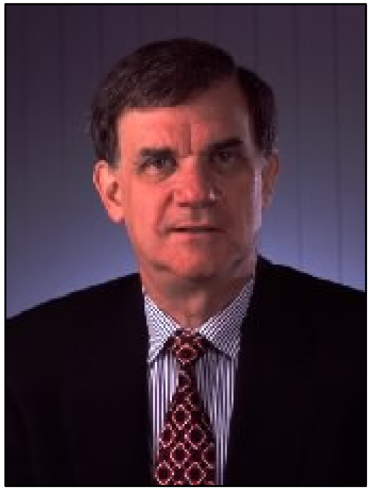
- but only with respect to transfer into practice (there are other measures of research success)

Administration

An initiative of ACM SIGSOFT

- volunteers mostly pay their own way
- modest funding from US NSF, UK IEE, and various agencies in Italy, UK, Germany, and Japan

International executive committee



Leon Osterweil

Carlo Ghezzi

Jeff Kramer

Alexander Wolf

Method

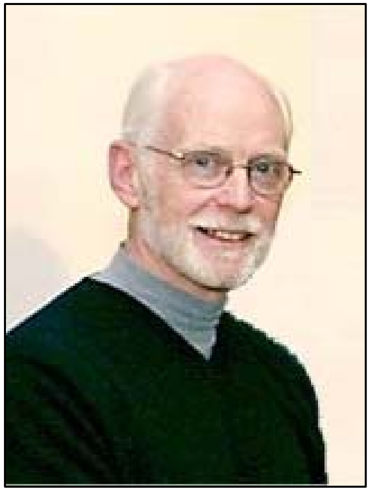
Form teams around practices

- important and widespread
  - » examples: configuration management, programming languages, middleware, assertions, walkthroughs, …
- recruit volunteers, including researchers (academic and industrial) and practitioners

Start from practice and trace backward

Use accepted historical tools

- qualify conclusions by solidity of evidence; use references, oral histories, …



Michael Mahoney

Professor of (Science) History

A quick tour of two completed reports

Software configuration management

- first report to emerge from the project
- ACM TOSEM, October 2005

Middleware technology

- latest report to emerge from the project
- ACM TOSEM, to appear October 2008

Software configuration management

Investigation into the research origins of successful SCM vendor products (ca. 2003)

- version control
- product models
- change control
- composition/selection
- build management
- workspace management

Lead authors

Jacky Estublier David Leblang

Other team members

- G. Clemm, IBM (ClearCase)
- R. Conradi, U. Trondheim
- A. van der Hoek, U. California
- W. Tichy, U. Karlsruhe (RCS)
- D. Wiborg-Weber, Telelogic (Continuus)

Is there a practice?

Practice measured by sales of vendor products

- Ovum: $1B (1998), $2B (2000), $3.3B (2002)
  - » 25% mainframe
  - » 15%-20% workstations
  - » 5%-10% PC
- Gartner: $6B (2003)

BTW: this ignores the “sales” of freeware and shareware

- examples: CVS, Subversion

Historical method

1. Examine characteristics/features of leading products in SCM market

2. Assume products are used in practice

3. Trace product characteristics/features back to research ideas and prototypes

4. Make arguments for/against influence of research on practice via products

When was it introduced?

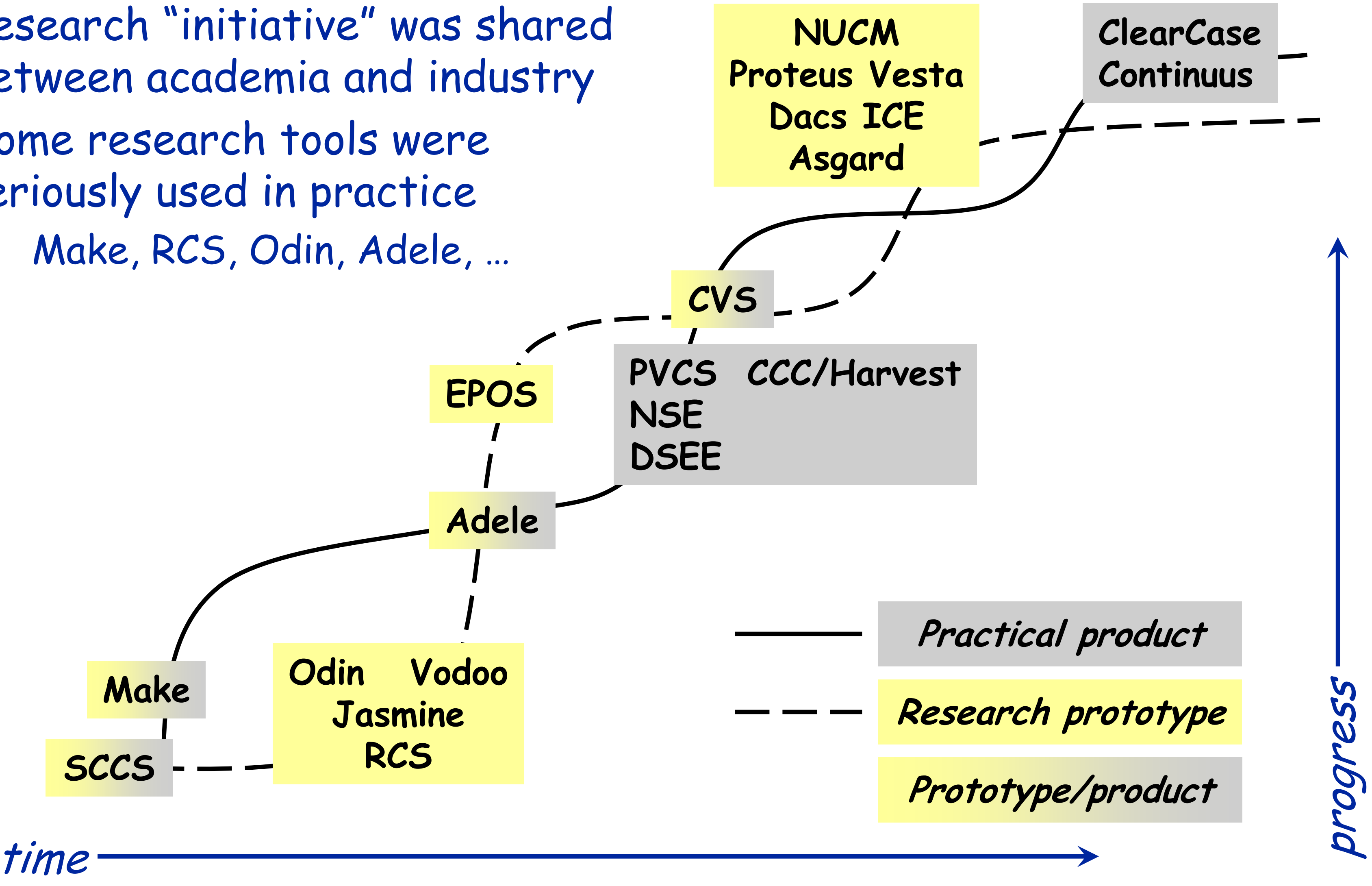
<table>
<tr><th></th><th><em>Academic Research</em></th><th><em>Industrial Research</em></th><th><em>Industrial Product</em></th></tr>
<tr><td><strong>1972</strong></td><td></td><td>SCCS (Bell Labs)</td><td></td></tr>
<tr><td><strong>1976</strong></td><td></td><td>Diff (Bell Labs)</td><td></td></tr>
<tr><td><strong>1977</strong></td><td></td><td>Make (Bell Labs)</td><td></td></tr>
<tr><td><strong>1980</strong></td><td>Variants, RCS (Purdue)</td><td>Change-sets (Xerox Parc)</td><td></td></tr>
<tr><td><strong>1982</strong></td><td>Merging, and/or graph (Purdue)</td><td></td><td></td></tr>
<tr><td><strong>1983</strong></td><td></td><td colspan="2">Change-sets (Aide-de-Camp)</td></tr>
<tr><td><strong>1984</strong></td><td>Selection (Grenoble)</td><td></td><td></td></tr>
<tr><td><strong>1985</strong></td><td></td><td colspan="2">System model (DSEE)</td></tr>
<tr><td rowspan="2"><strong>1988</strong></td><td colspan="2">Process support (Grenoble)</td><td></td></tr>
<tr><td colspan="3">NSE Workspaces (Carnegie Mellon, Sun)</td></tr>
<tr><td><strong>1990</strong></td><td></td><td>nDFS file system (Bell Labs)</td><td></td></tr>
<tr><td><strong>1994</strong></td><td></td><td colspan="2">Virtual file system and MultiSite (ClearCase)</td></tr>
<tr><td><strong>1996</strong></td><td></td><td>Activities (Asgard, Bellcore)</td><td></td></tr>
<tr><td><strong>2000</strong></td><td colspan="2">WebDAV (California, Microsoft, ClearCase, …)</td><td></td></tr>
</table>

An argument: research/product timing

Research "initiative" was shared between academia and industry

Some research tools were seriously used in practice

- Make, RCS, Odin, Adele, …



An argument: professional interaction

SCM research community organized regular workshops beginning in 1988

Product architects were present at all

- Cagan, Clemm, Dart, Leblang, Wiborg-Weber, ...
- Some presented work, while others simply attended and participated in discussion

The meetings put ideas “in the air”, and helped to keep them there

The role of creativity: vendor's view

Vendors tend to consider that research impact is restricted to…

- *algorithms* (e.g., differencing)
- *pieces of reusable code* (e.g., RCS)

and not…

- *concepts* (e.g., hierarchical workspaces)
- *architectures* (e.g., peer-to-peer repositories)

which are often seen as "**engineering common sense**"

The role of creativity: researcher's view

Researchers tend to consider that…

- *precedence*
- *concepts*
- *prototypes*

are sufficient as impact and ignore…

- *efficiency*
- *usability*
- *reliability*

dismissing them as "**engineering common sense**"

Both are right and both are wrong

A good idea is had more than once

Vendors have disincentives for distributing credit for ideas

Researchers have incentives for claiming credit for ideas

Research and productization both require ***engineering creativity***

Middleware technology

Investigation into the research origins of successful middleware technology (ca. 2007)

- web services
- application servers
- transaction monitors
- distributed object systems
- message queues
- remote procedure call systems

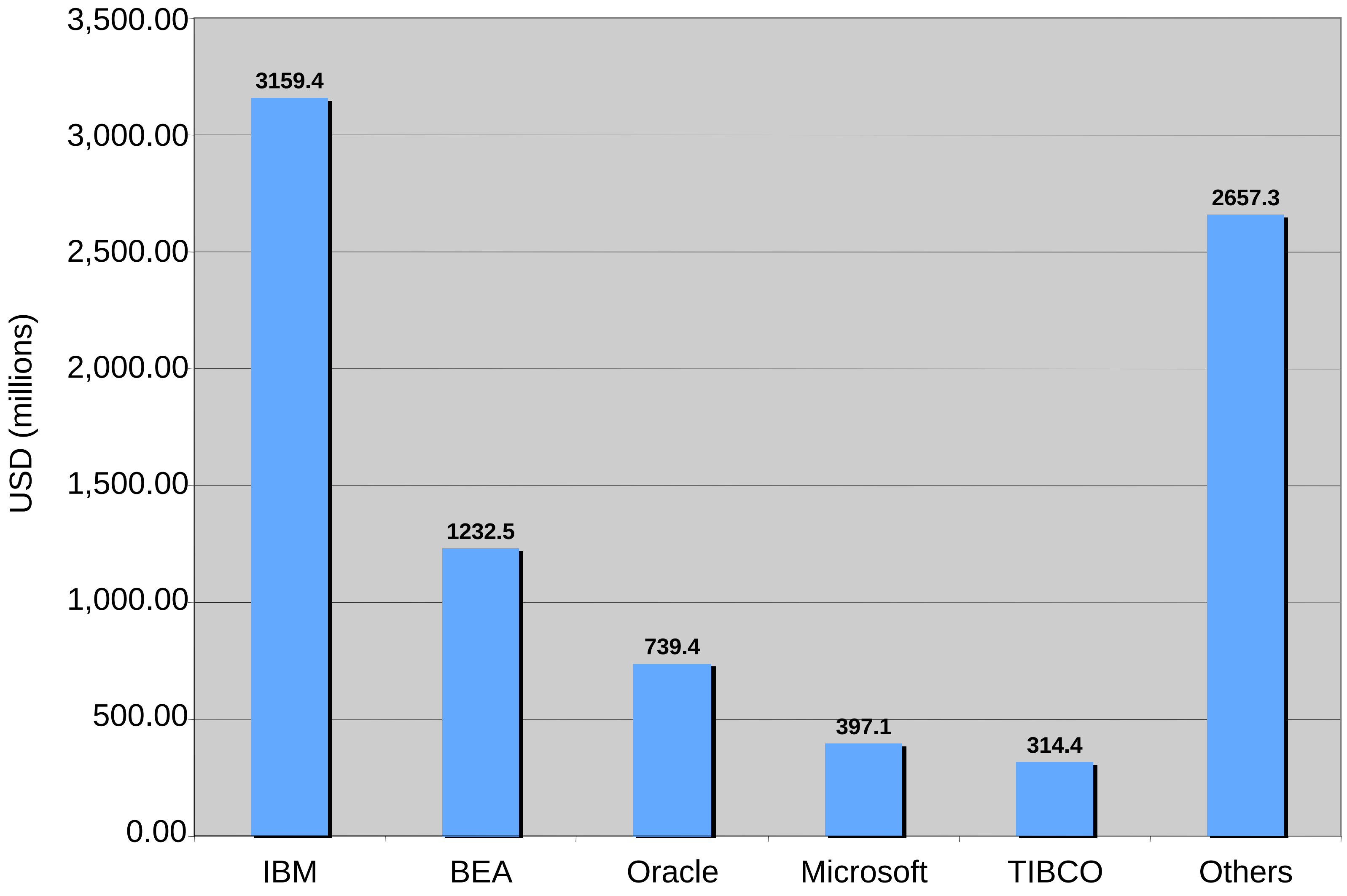
Lead authors

Wolfgang Emmerich

Mikio Aoyama

Joe Sventek

Is there a practice?

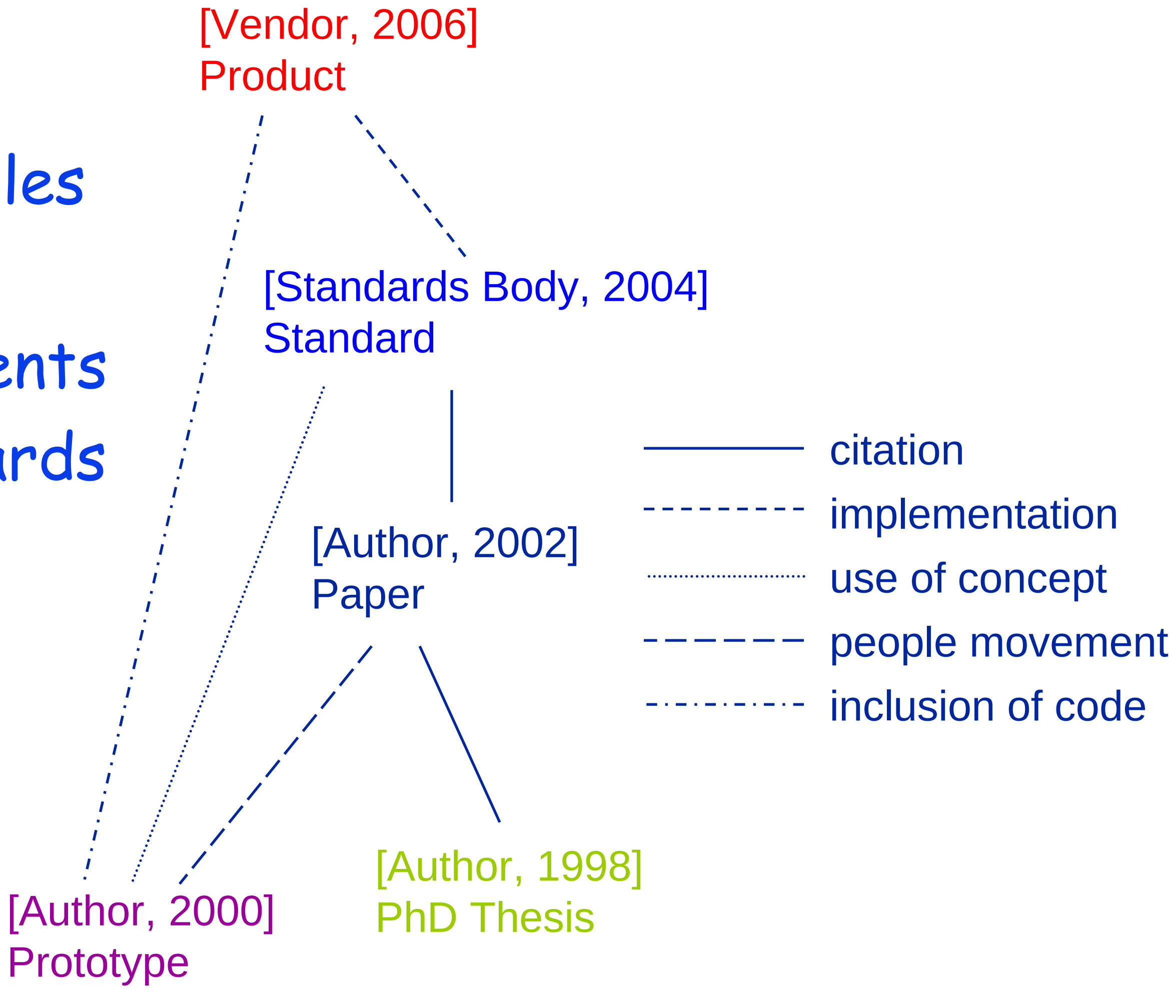


Middleware License Market in 2005 [Gartner 2006]

Historical method

1. Seek sources
   - market analysis reports
   - professional articles
   - technical reports
   - standards documents
   - minutes of standards meetings
   - people movement
   - PhD theses
   - software
   - interviews

2. Derive "impact trace graph"



[Vendor, 2006]
Product

[Standards Body, 2004]
Standard

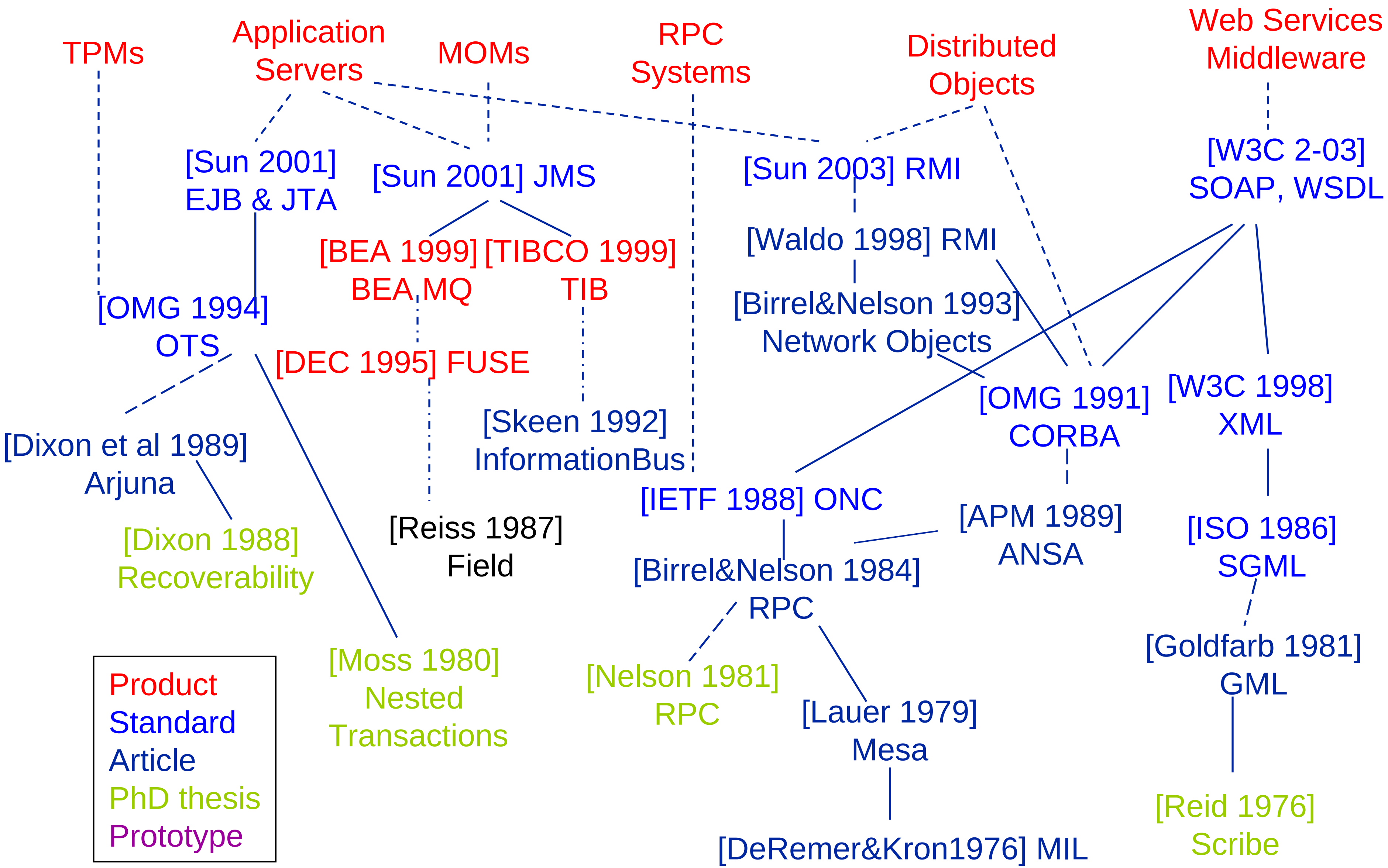
[Author, 2002]
Paper

[Author, 2000]
Prototype

[Author, 1998]
PhD Thesis

─── citation
- - - - implementation
········ use of concept
— — — people movement
-·-·-·- inclusion of code

The big picture of impact



TPMs

Application Servers

MOMs

RPC Systems

Distributed Objects

Web Services Middleware

[Sun 2001] EJB & JTA

[Sun 2001] JMS

[Sun 2003] RMI

[W3C 2-03] SOAP, WSDL

[BEA 1999] BEA MQ

[TIBCO 1999] TIB

[Waldo 1998] RMI

[OMG 1994] OTS

[Birrel&Nelson 1993] Network Objects

[DEC 1995] FUSE

[OMG 1991] CORBA

[W3C 1998] XML

[Skeen 1992] InformationBus

[Dixon et al 1989] Arjuna

[IETF 1988] ONC

[APM 1989] ANSA

[ISO 1986] SGML

[Dixon 1988] Recoverability

[Reiss 1987] Field

[Birrel&Nelson 1984] RPC

[Goldfarb 1981] GML

[Moss 1980] Nested Transactions

[Nelson 1981] RPC

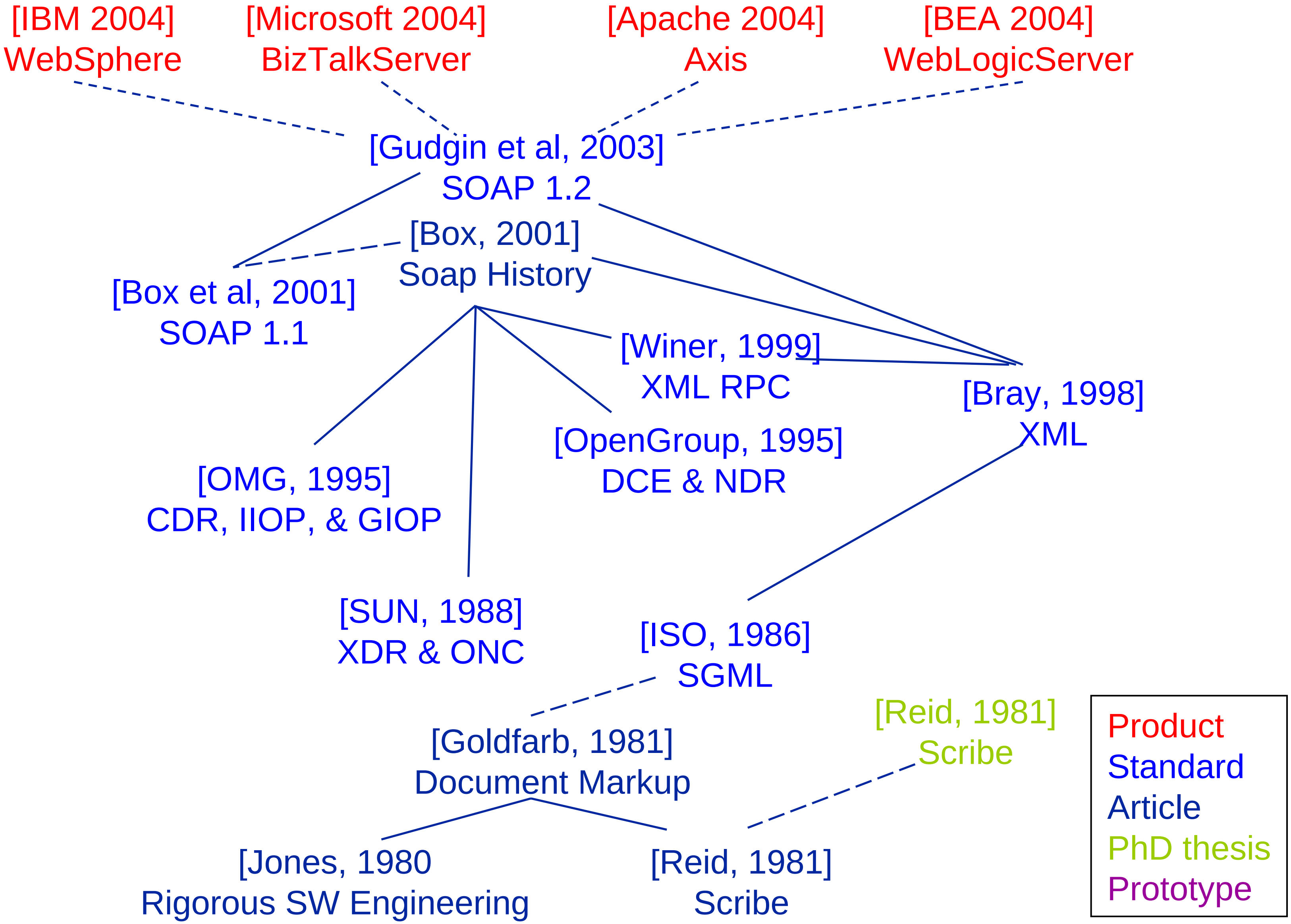
[Lauer 1979] Mesa

[Reid 1976] Scribe

[DeRemer&Kron1976] MIL

Product
Standard
Article
PhD thesis
Prototype

Trace: Simple Object Access Protocol



[IBM 2004]
WebSphere

[Microsoft 2004]
BizTalkServer

[Apache 2004]
Axis

[BEA 2004]
WebLogicServer

[Gudgin et al, 2003]
SOAP 1.2

[Box, 2001]
Soap History

[Box et al, 2001]
SOAP 1.1

[Winer, 1999]
XML RPC

[Bray, 1998]
XML

[OpenGroup, 1995]
DCE & NDR

[OMG, 1995]
CDR, IIOP, & GIOP

[SUN, 1988]
XDR & ONC

[ISO, 1986]
SGML

[Reid, 1981]
Scribe

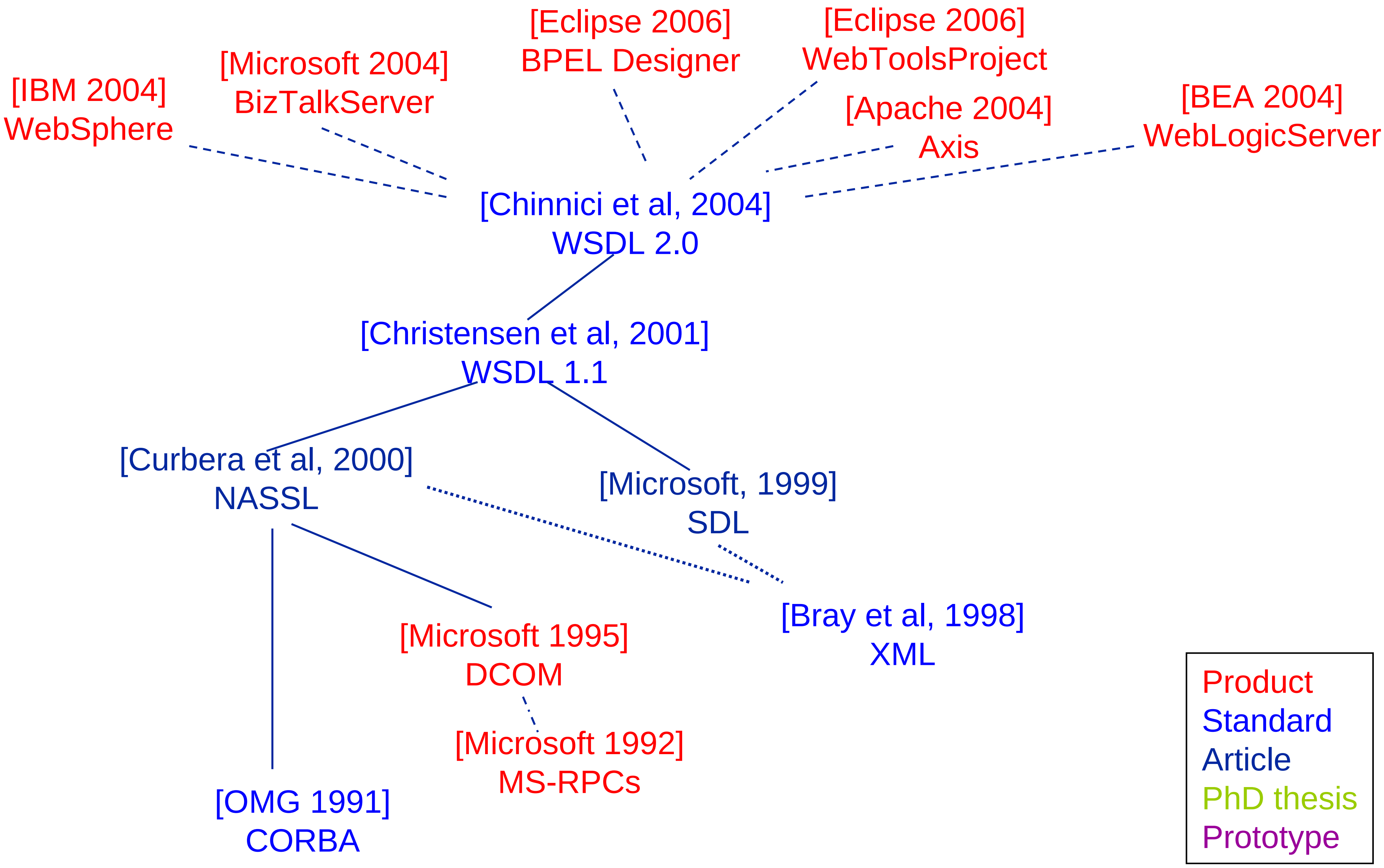
[Goldfarb, 1981]
Document Markup

[Jones, 1980
Rigorous SW Engineering

[Reid, 1981]
Scribe

Product
Standard
Article
PhD thesis
Prototype

Trace: Web Services Description Lang.



[Eclipse 2006]
BPEL Designer

[Eclipse 2006]
WebToolsProject

[Microsoft 2004]
BizTalkServer

[IBM 2004]
WebSphere

[Apache 2004]
Axis

[BEA 2004]
WebLogicServer

[Chinnici et al, 2004]
WSDL 2.0

[Christensen et al, 2001]
WSDL 1.1

[Curbera et al, 2000]
NASSL

[Microsoft, 1999]
SDL

[Bray et al, 1998]
XML

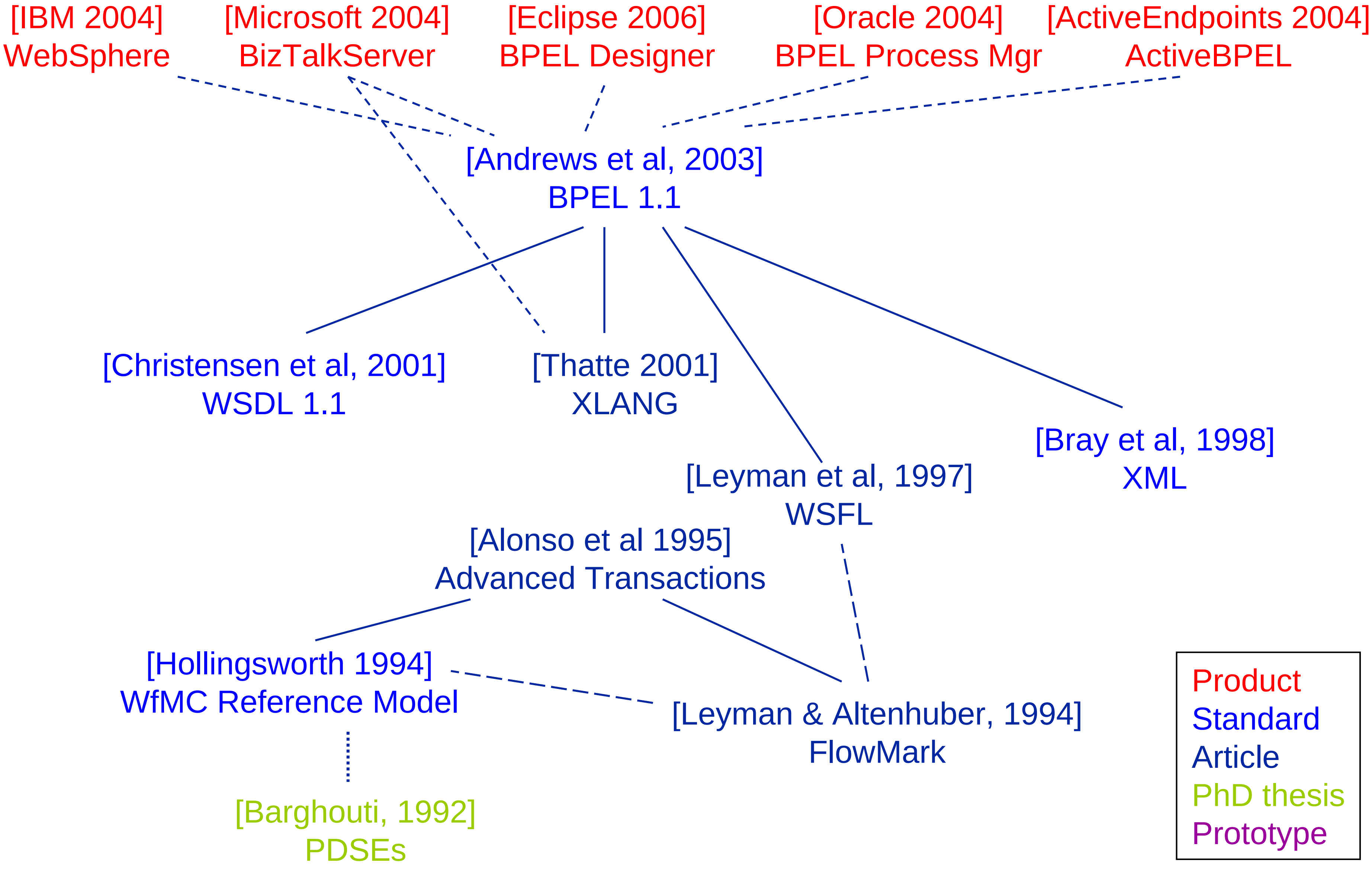
[Microsoft 1995]
DCOM

[Microsoft 1992]
MS-RPCs

[OMG 1991]
CORBA

Product
Standard
Article
PhD thesis
Prototype

Trace: Business Process Execution Lang.



[IBM 2004]
WebSphere

[Microsoft 2004]
BizTalkServer

[Eclipse 2006]
BPEL Designer

[Oracle 2004]
BPEL Process Mgr

[ActiveEndpoints 2004]
ActiveBPEL

[Andrews et al, 2003]
BPEL 1.1

[Christensen et al, 2001]
WSDL 1.1

[Thatte 2001]
XLANG

[Bray et al, 1998]
XML

[Leyman et al, 1997]
WSFL

[Alonso et al 1995]
Advanced Transactions

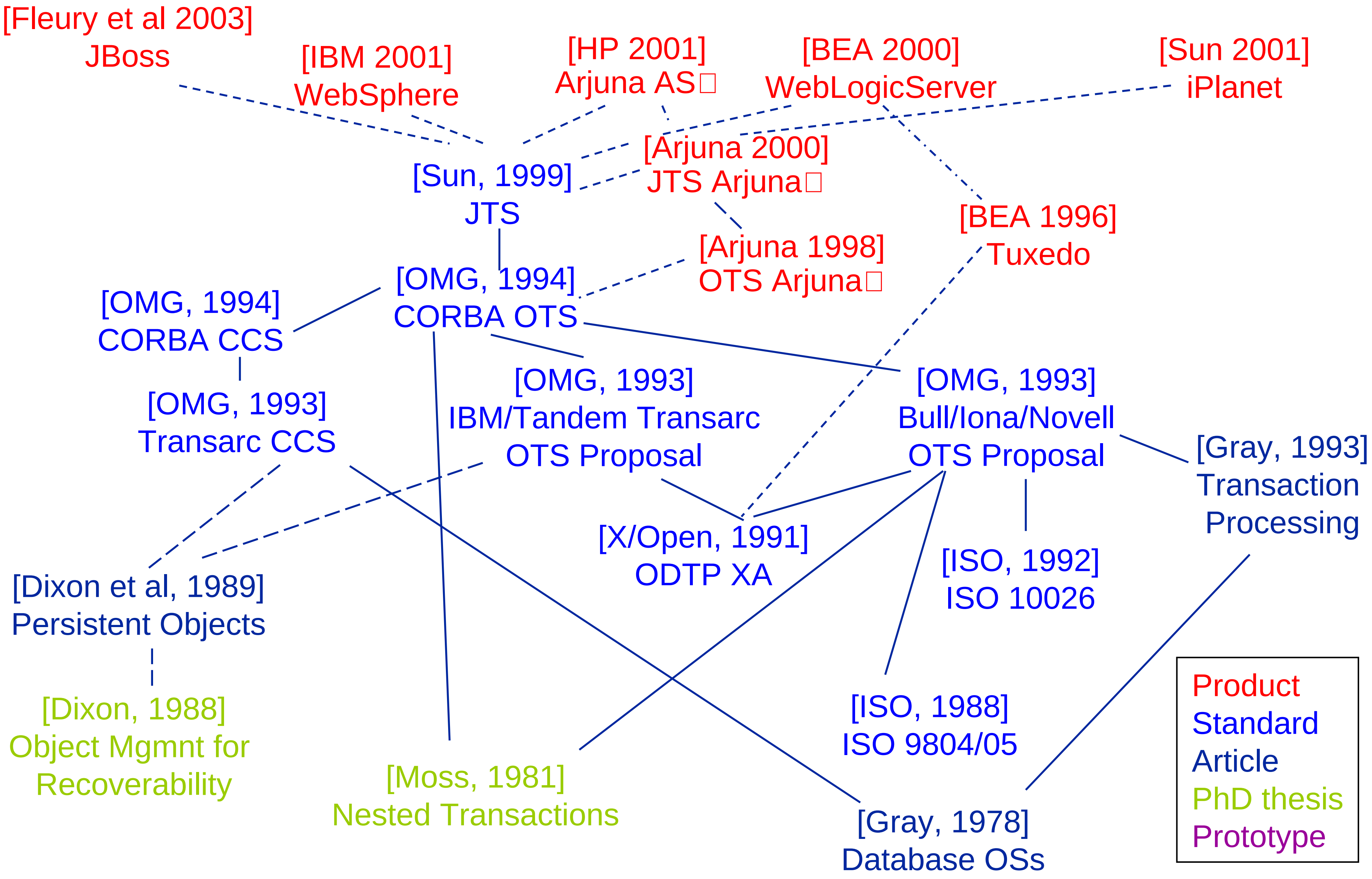
[Hollingsworth 1994]
WfMC Reference Model

[Leyman & Altenhuber, 1994]
FlowMark

[Barghouti, 1992]
PDSEs

Product
Standard
Article
PhD thesis
Prototype

Trace: Transactions in App. Servers



[Fleury et al 2003]
JBoss

[IBM 2001]
WebSphere

[HP 2001]
Arjuna AS□

[BEA 2000]
WebLogicServer

[Sun 2001]
iPlanet

[Arjuna 2000]
JTS Arjuna□

[Sun, 1999]
JTS

[BEA 1996]
Tuxedo

[Arjuna 1998]
OTS Arjuna□

[OMG, 1994]
CORBA OTS

[OMG, 1994]
CORBA CCS

[OMG, 1993]
Transarc CCS

[OMG, 1993]
IBM/Tandem Transarc
OTS Proposal

[OMG, 1993]
Bull/Iona/Novell
OTS Proposal

[Gray, 1993]
Transaction
Processing

[X/Open, 1991]
ODTP XA

[ISO, 1992]
ISO 10026

[Dixon et al, 1989]
Persistent Objects

[Dixon, 1988]
Object Mgmnt for
Recoverability

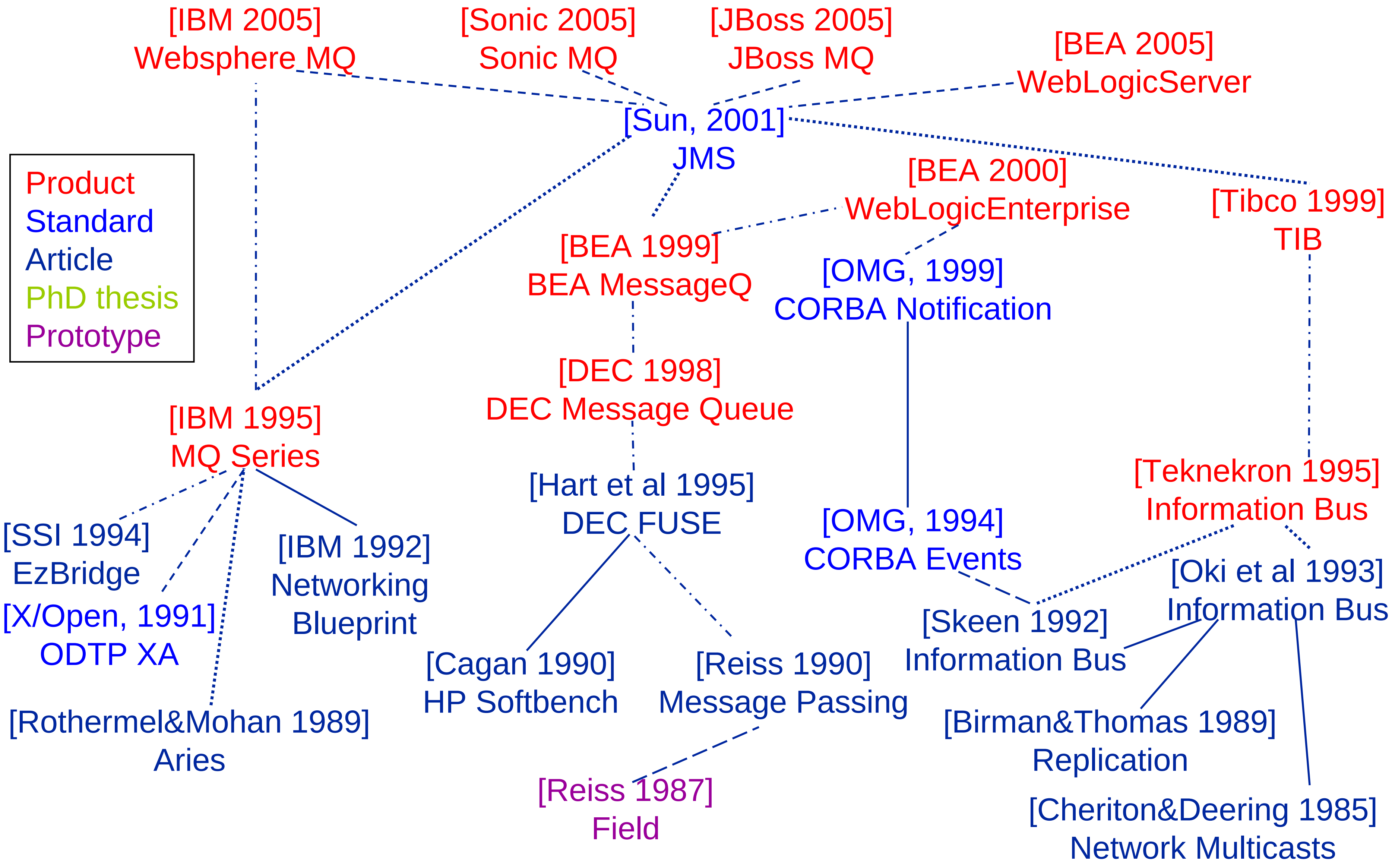
[Moss, 1981]
Nested Transactions

[ISO, 1988]
ISO 9804/05

[Gray, 1978]
Database OSs

Product
Standard
Article
PhD thesis
Prototype

Trace: Messaging in App. Servers



Product
Standard
Article
PhD thesis
Prototype

[IBM 2005]
Websphere MQ

[Sonic 2005]
Sonic MQ

[JBoss 2005]
JBoss MQ

[BEA 2005]
WebLogicServer

[Sun, 2001]
JMS

[BEA 2000]
WebLogicEnterprise

[Tibco 1999]
TIB

[BEA 1999]
BEA MessageQ

[OMG, 1999]
CORBA Notification

[DEC 1998]
DEC Message Queue

[IBM 1995]
MQ Series

[Hart et al 1995]
DEC FUSE

[Teknekron 1995]
Information Bus

[OMG, 1994]
CORBA Events

[SSI 1994]
EzBridge

[Oki et al 1993]
Information Bus

[IBM 1992]
Networking
Blueprint

[Skeen 1992]
Information Bus

[X/Open, 1991]
ODTP XA

[Cagan 1990]
HP Softbench

[Reiss 1990]
Message Passing

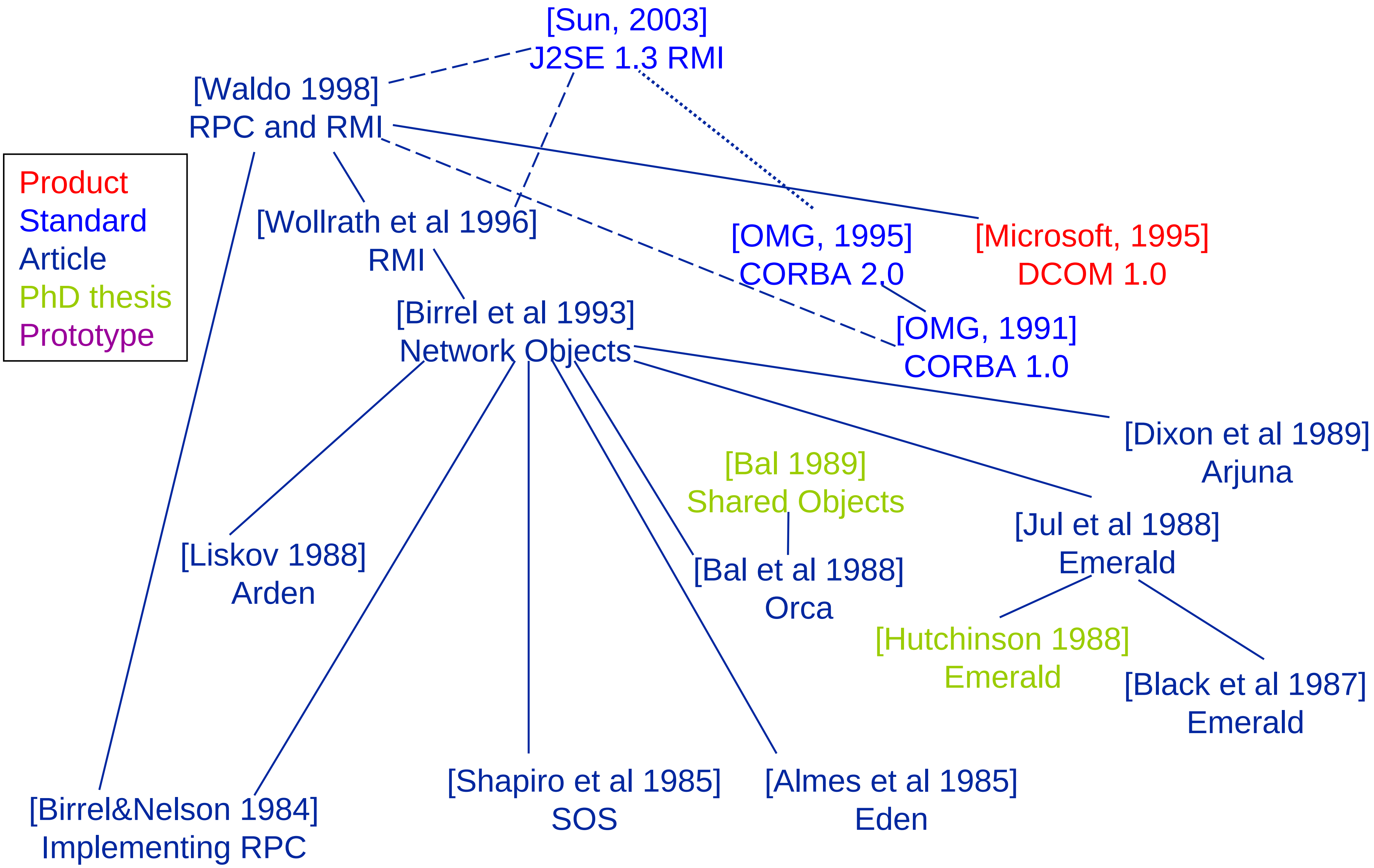
[Rothermel&Mohan 1989]
Aries

[Birman&Thomas 1989]
Replication

[Reiss 1987]
Field

[Cheriton&Deering 1985]
Network Multicasts

Trace: Dist. Objects in App. Servers



Product
Standard
Article
PhD thesis
Prototype

[Sun, 2003]
J2SE 1.3 RMI

[Waldo 1998]
RPC and RMI

[Wollrath et al 1996]
RMI

[OMG, 1995]
CORBA 2.0

[Microsoft, 1995]
DCOM 1.0

[Birrel et al 1993]
Network Objects

[OMG, 1991]
CORBA 1.0

[Dixon et al 1989]
Arjuna

[Bal 1989]
Shared Objects

[Jul et al 1988]
Emerald

[Liskov 1988]
Arden

[Bal et al 1988]
Orca

[Hutchinson 1988]
Emerald

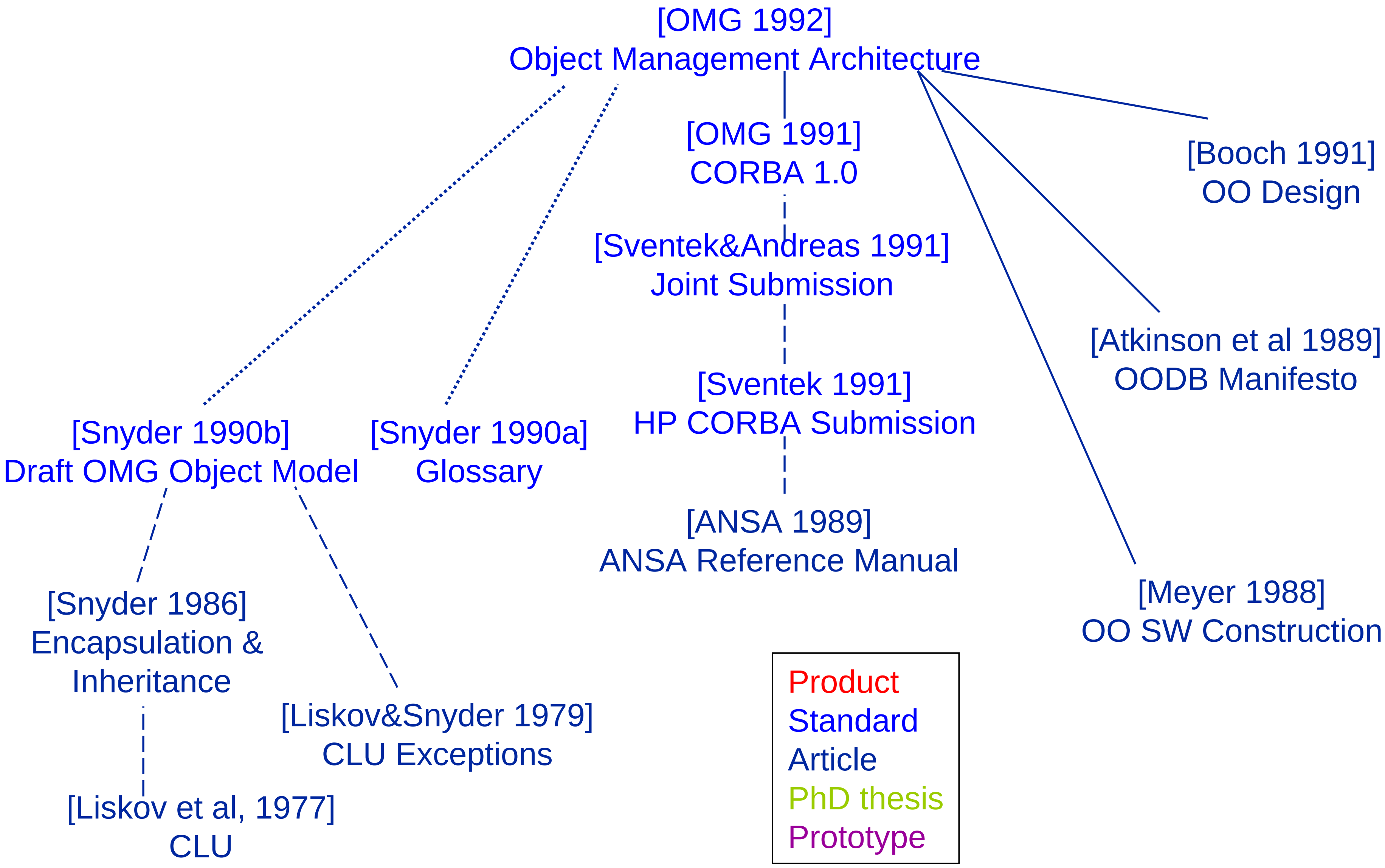
[Black et al 1987]
Emerald

[Shapiro et al 1985]
SOS

[Almes et al 1985]
Eden

[Birrel&Nelson 1984]
Implementing RPC

Trace: Distributed Objects in CORBA



[OMG 1992]
Object Management Architecture

[OMG 1991]
CORBA 1.0

[Sventek&Andreas 1991]
Joint Submission

[Sventek 1991]
HP CORBA Submission

[ANSA 1989]
ANSA Reference Manual

[Booch 1991]
OO Design

[Atkinson et al 1989]
OODB Manifesto

[Meyer 1988]
OO SW Construction

[Snyder 1990b]
Draft OMG Object Model

[Snyder 1990a]
Glossary

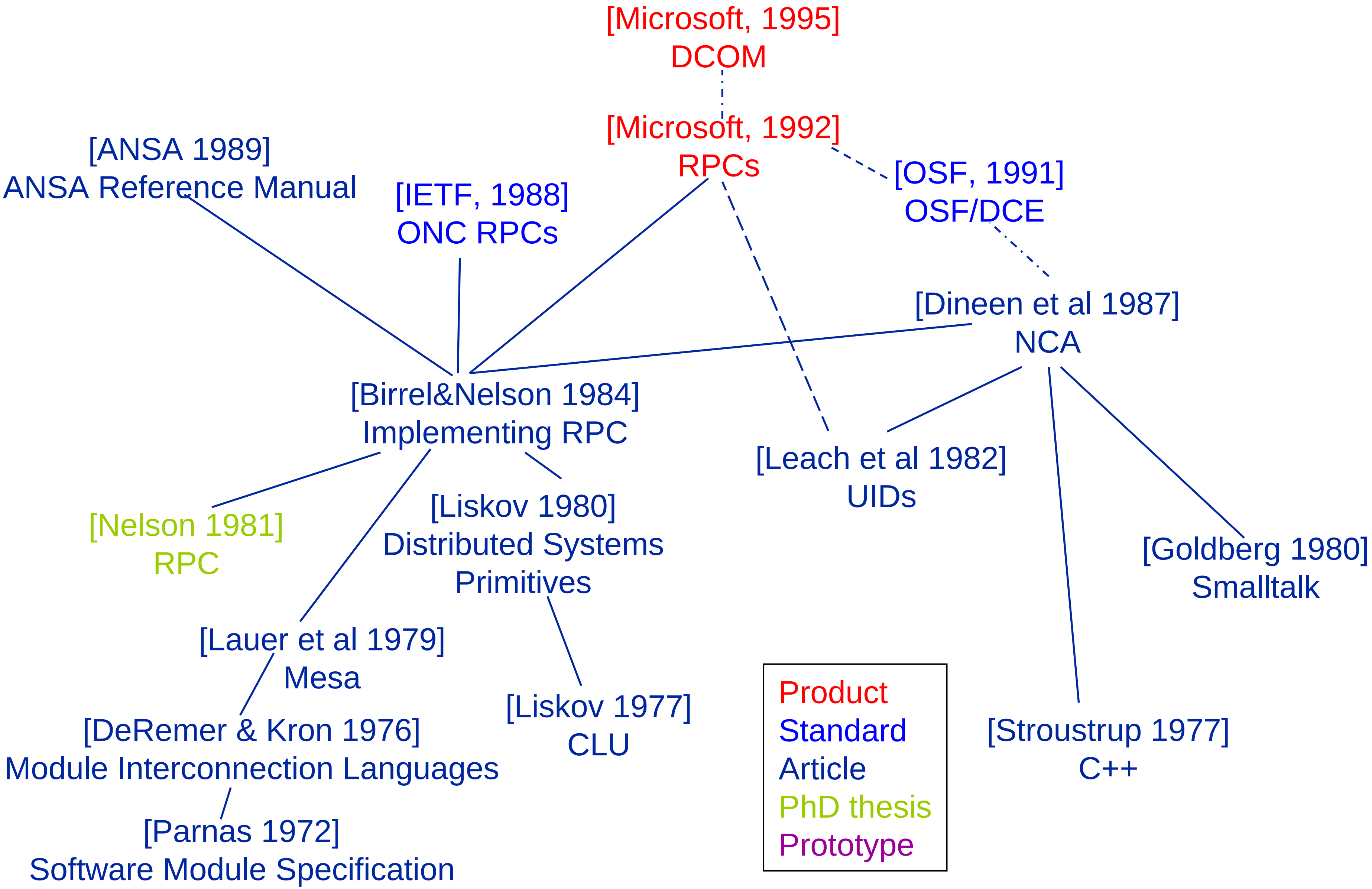
[Snyder 1986]
Encapsulation &
Inheritance

[Liskov&Snyder 1979]
CLU Exceptions

[Liskov et al, 1977]
CLU

Product
Standard
Article
PhD thesis
Prototype

Trace: Remote Procedure Calls



[Microsoft, 1995]
DCOM

[Microsoft, 1992]
RPCs

[ANSA 1989]
ANSA Reference Manual

[IETF, 1988]
ONC RPCs

[OSF, 1991]
OSF/DCE

[Dineen et al 1987]
NCA

[Birrel&Nelson 1984]
Implementing RPC

[Leach et al 1982]
UIDs

[Nelson 1981]
RPC

[Liskov 1980]
Distributed Systems
Primitives

[Goldberg 1980]
Smalltalk

[Lauer et al 1979]
Mesa

[Liskov 1977]
CLU

[Stroustrup 1977]
C++

[DeRemer & Kron 1976]
Module Interconnection Languages

[Parnas 1972]
Software Module Specification

Product
Standard
Article
PhD thesis
Prototype

M/W: Some key findings/confirmations

Technology development is interdisciplinary

- often winds back and forth among disciplines

Technology maturation needs time

- 15-20 years between first publication of an idea and widespread availability in products

Technology transfer needs commitment

- people movement is most successful vehicle

PhD students are critical sources of ideas

- almost all impact traces lead back to PhD theses

Standards are critical enablers of ideas

- without widespread agreements on ideas there is no widespread adoption

Tech. development is interdisciplinary

Impact traces often cross CS disciplines

For middleware…

- software engineering
- networking
- programming languages
- distributed systems
- databases

Impact sometimes larger in area other than first publication

- e.g., message queues

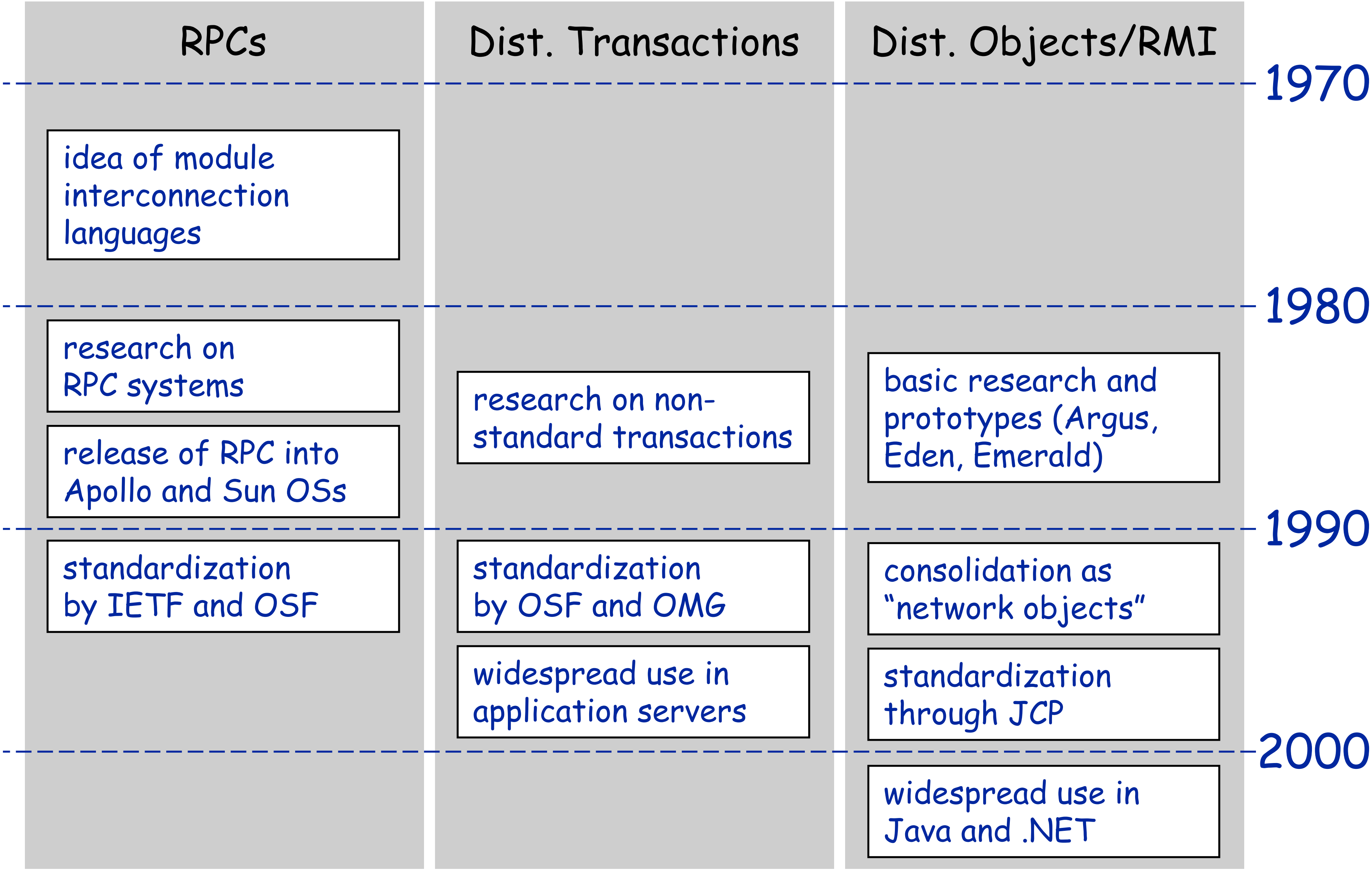
Example: RPC IDLs

- infor. hiding [CACM 15(5), 1972]
- MIL [IEEE TSE SE-2(2), 1976]
- Mesa [ICSE-4, 1977]
- Cedar RPCs [ACM ToCS 2(1), 1984]
- Sun RPC [IETF RFC 1057, 1987]

Example: dist. transactions

- operating systems (Gray, 1976)
- nested transactions (Moss, 1981)
- concur. control (Bernstein, 1987)
- Arjuna (Dixon, 1989)
- OSF ODTP/XA (1991)
- CORBA CCS, OTS (1994)
- J2EE JTS, JTA (2001)

Technology maturation needs time



	RPCs	Dist. Transactions	Dist. Objects/RMI
1970	idea of module interconnection languages		
1980	research on RPC systems release of RPC into Apollo and Sun OSs	research on non-standard transactions	basic research and prototypes (Argus, Eden, Emerald)
1990	standardization by IETF and OSF	standardization by OSF and OMG widespread use in application servers	consolidation as "network objects" standardization through JCP
2000			widespread use in Java and .NET

Technology transfer needs commitment

- B. Nelson from CMU to Xerox PARC
  - wrote definitive paper on RPCs with A. Birrel
- B. Nelson and A. Birrel to DEC Research
  - wrote Network Object paper providing the basis for Java RMI
- A. Watson from APM to OMG
  - lead CORBA standardization
- A. Herbert from Cambridge to APM
  - devised ANSA
- G. Dixon from NCL to Transarc
  - wrote OMG CORBA OTS and CCS service specs
- J. Waldo from UMass, J. Sventek from APM to HP
  - wrote CORBA 1.0 spec
- J. Waldo from HP to Sun
  - wrote RMI spec

PhD students are critical sources

Remote procedure calls

- architecture and failure semantics: Nelson (CMU 1981)
- orphan detection: Panzieri (Newcastle 1985)

Distributed transactions

- nested transactions: Moss (MIT 1981)
- object transactions: Dixon (Newcastle 1987)

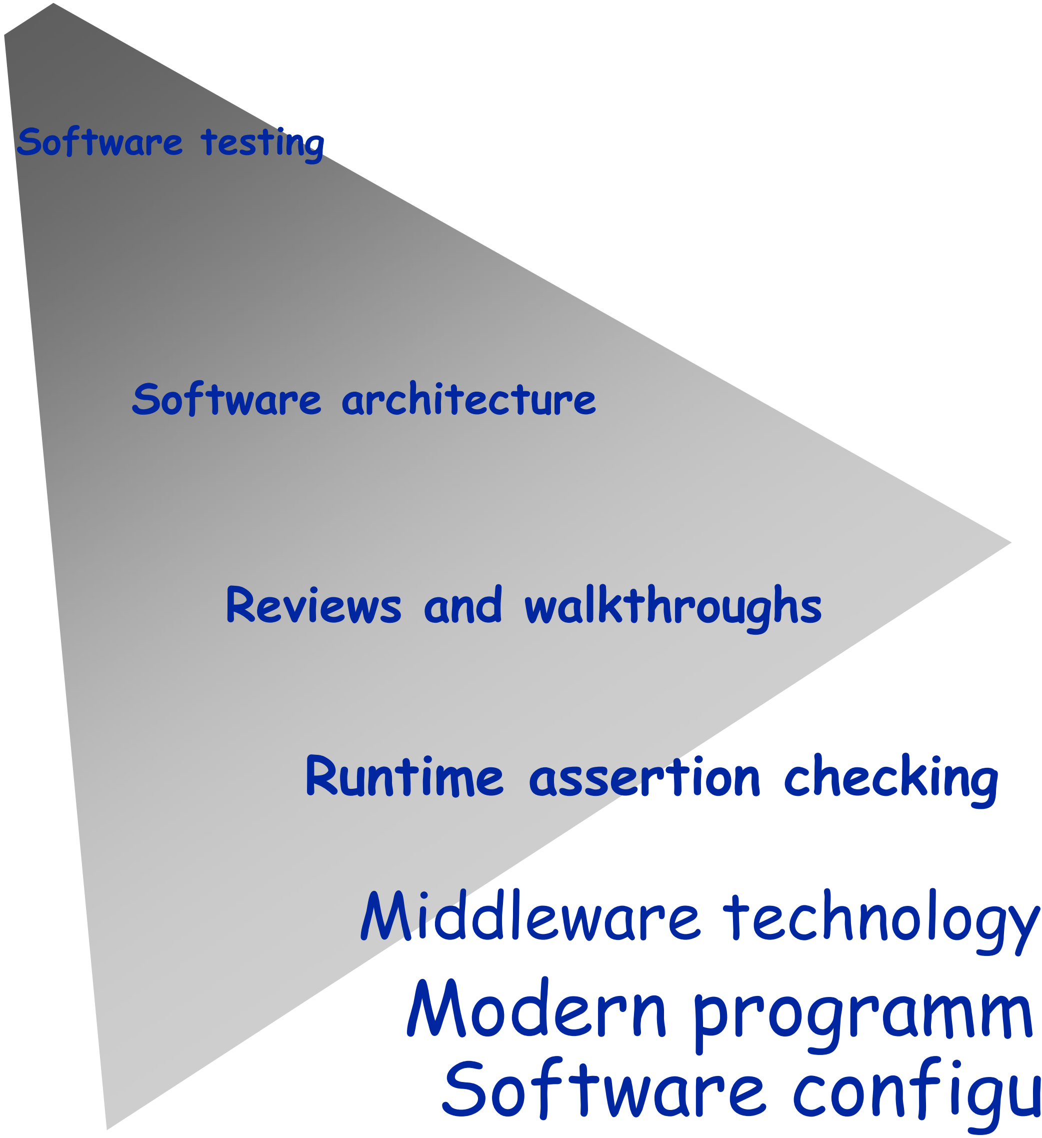
Distributed object models

- general object models: Snyder (MIT 1978)
- RMI object model: Hutchinson (UW 1987) and Bal (Vrije 1989)

Web services

- Scribe: Reid (CMU 1981)

Impact reports roadmap



???

Software testing

Software architecture

Reviews and walkthroughs

Runtime assertion checking

Middleware technology

Modern programming languages

Software configuration management

Preliminary project “meta“ findings

1. SE research has had impact on SE practice
2. Lasting impact comes most readily from repeated and sustained interactions
3. Interplay can be difficult to determine precisely and communicate clearly
4. Substantially different mechanisms have been successful at causing impact
5. More benefit from nurturing many and varied ways than single approach to impact
6. Community needs support to maintain the nurturing environment

Further information and reading

Impact web site

- http://www.acm.org/sigsoft/impact/
- contains links to published reports on
  - » software configuration management (ACM TOSEM)
  - » modern programming languages (ACM TOSEM)
  - » middleware technology (ACM TOSEM)
  - » run-time assertion checking (ACM SIGSOFT SEN)

Overview article

- to appear in IEEE Computer, early 2008