

SUMEX

**STANFORD UNIVERSITY
MEDICAL EXPERIMENTAL COMPUTER RESOURCE**

RR-00785

COMPETING RENEWAL APPLICATION

Submitted to

**BIOMEDICAL RESEARCH TECHNOLOGY PROGRAM
NATIONAL INSTITUTES OF HEALTH**

June 1, 1985

**STANFORD UNIVERSITY SCHOOL OF MEDICINE
Edward H. Shortliffe, Principal Investigator
Edward A. Feigenbaum, Co-Principal Investigator**

DEPARTMENT OF HEALTH AND HUMAN SERVICES
PUBLIC HEALTH SERVICE

GRANT APPLICATION

FOLLOW INSTRUCTIONS CAREFULLY

LEAVE BLANK

TYPE	ACTIVITY	NUMBER
REVIEW GROUP		FORMERLY
COUNCIL/BOARD (Month, year)		DATE RECEIVED

1. TITLE OF APPLICATION (Do not exceed 56 typewriter spaces)

SU Medical Experimental Computer Resource (SUMEX)

2. RESPONSE TO SPECIFIC PROGRAM ANNOUNCEMENT NO YES (If "YES," state RFA number and/or announcement title)

3. PRINCIPAL INVESTIGATOR/PROGRAM DIRECTOR

3a. NAME (Last, first, middle) Shortliffe, Edward H. 3b. SOCIAL SECURITY NUMBER [REDACTED]

3c. POSITION TITLE Associate Professor of Medicine 3d. MAILING ADDRESS (Street, city, state, zip code) Stanford University School of Medicine Room TB-105 Stanford, CA 94305

3e. DEPARTMENT, SERVICE, LABORATORY OR EQUIVALENT Department of Medicine

3f. MAJOR SUBDIVISION School of Medicine 3g. TELEPHONE (Area code, number and extension) (415)497-6979

4. HUMAN SUBJECTS NO YES Exemption # _____ OR Form HHS 596 enclosed 5. RECOMBINANT DNA NO YES

6. DATES OF ENTIRE PROPOSED PROJECT PERIOD From: 8/1/86 Through: 7/31/91 7. DIRECT COSTS REQUESTED FOR FIRST 12-MONTH BUDGET PERIOD (from page 4) \$ 1,370,257 8. DIRECT COSTS REQUESTED FOR ENTIRE PROPOSED PROJECT PERIOD (from page 5) \$ 6,963,247

9. PERFORMANCE SITES (Organizations and addresses) Stanford University School of Medicine Stanford, CA 94305 10. INVENTIONS (Competing continuation application only) NO YES Previously reported OR Not previously reported

11. APPLICANT ORGANIZATION (Name, address, and congressional district) Stanford University c/o Sponsored Projects Office Encina Hall, Room 40 Stanford, CA 94305 Congressional District No. 11

12. TYPE OF ORGANIZATION Public. Specify Federal State Local Private Nonprofit For Profit (General) For Profit (Small Business) 13. ENTITY IDENTIFICATION NUMBER IRS No. 94-1156365

14. ORGANIZATIONAL COMPONENT TO RECEIVE CREDIT FOR BIOMEDICAL RESEARCH SUPPORT GRANT Code Description

15. OFFICIAL IN BUSINESS OFFICE TO BE NOTIFIED IF AN AWARD IS MADE (Name, title, address and telephone number.) Patricia Byers, Contract Officer Sponsored Projects Office Encina Hall, Room 40 Stanford, CA 94305 (415)497-2883 16. OFFICIAL SIGNING FOR APPLICANT ORGANIZATION (Name, title, address and telephone number) PATRICIA BYERS SENIOR CONTRACT OFFICER Sponsored Projects Office Encina Hall, Room 40 Stanford, CA 94305 (415)497-2883

17. PRINCIPAL INVESTIGATOR/PROGRAM DIRECTOR ASSURANCE: I agree to accept responsibility for the scientific conduct of the project and to provide the required progress reports if a grant is awarded as a result of this application. Willful provision of false information is a criminal offense (U.S. Code, Title 18, Section 1001). SIGNATURE OF PERSON NAMED IN 3a (In ink. "Per" signature not acceptable) Edward H Shortliffe DATE 5/24/85

18. CERTIFICATION AND ACCEPTANCE: I certify that the statements herein are true and complete to the best of my knowledge, and accept the obligation to comply with Public Health Service terms and conditions if a grant is awarded as the result of this application. A willfully false certification is a criminal offense (U.S. Code, Title 18, Section 1001). SIGNATURE OF PERSON NAMED IN 16 (In ink. "Per" signature not acceptable) Patricia Byers DATE 5/29/85

ABSTRACT OF RESEARCH PLAN

KEY PROFESSIONAL EXPERIENCE AND EDUCATION OF PI

NAME	POSITION	DEGREE AND INSTITUTION
E. Shortliffe	Principal Investigator	Medicine & Computer Science
E. Feigenbaum	Co-Principal Investigator	Computer Science
T. Rindfleisch	Director of Resource	Medicine & Computer Science
C. Jacobs	ONCOCIN Investigator	Medicine
L. Fagan	AIM Liaison	Medicine
W. Yeager	Systems Programmer/Assistant Resource Director	Medicine
B. Buchanan	Professor of Research - Computer Science	Computer Science
B. Hayes-Roth	Senior Research Associate	Computer Science
H. Brown	Senior Research Associate	Computer Science
P. Nii	Research Associate	Computer Science

ABSTRACT OF RESEARCH PLAN: State the application's long-term objectives and specific aims, making reference to the breadth and depth of the project, and describe concisely the methodology for achieving these goals. Avoid summaries of past accomplishments and the use of the first person. The abstract is meant to serve as a succinct and accurate description of the proposed work when separated from the application. **DO NOT EXCEED THE SPACE PROVIDED.**

Stanford University is developing and operating a national shared computing resource (SUMEX-AIM), in partnership with the NIH Biomedical Research Technology Program, to explore applications of computer science research in artificial intelligence (AI) to health research. There are three main objectives of the resource: 1) to develop and provide the computing resources and human assistance needed by scientists working on a broad range of biomedical applications of AI; 2) to demonstrate that it is feasible to provide resources and assistance to a national community of researchers, integrating distributed and centralized computing technology with local and national computer communication networks; and 3) to develop the community of scientists interested in working on AI in Medicine (AIM), promoting its growth and vigor through electronic communications. Besides the economic advantages of resource sharing made possible by electronic communications, we believe that a new style of science is emerging from communications-enhanced settings.

AI research is aimed at understanding the principles of computer-based symbolic knowledge representation, reasoning, and problem-solving processes and applying these to increase the computer's effectiveness as a tool in knowledge-intensive fields like medicine and biology. Our research work is driven by real-world scientific applications, chosen because of their relevance to current important biomedical problems and because they expose key underlying AI research issues. Current application areas include programs for differential diagnosis, cancer chemotherapy protocol management, protein structure inference, and drug interaction advice. Resource core research goals include basic research in areas such as blackboard problem-solving architectures and knowledge acquisition; methodologies for clinical decision-making advisors; and the development of network-based Lisp workstation computing environments.

Additional resource users will be selected within available resource capacity with the help of an AIM Executive Committee and Advisory Group on the basis of reviews of the proposed research. Selection criteria will include general scientific interest and merit, relevance to the resource AI mission, and the community orientation of the collaborator.

Table of Contents

1. Budget, Biographies, and Environment	1
1.1. Total Resource Budget	1
1.1.1. Budget for First 12-Month Period	2
1.1.2. Budget for Entire 5-Year Project Period	4
1.2. 2060 Operations Budget	5
1.2.1. Budget for First 12-Month Period	6
1.2.2. Budget for Entire 5-Year Project Period	8
1.3. Budget Explanation and Justification	9
1.3.1. Total Resource Budget	9
1.3.2. 2060 Operations Budget	14
1.4. Biographical Sketches	16
1.5. Current and Pending Support	36
1.6. Resources and Environment	51
2. Resource Plan	53
2.1. Introduction and Background	54
2.1.1. Principal Investigators' Executive Summary	54
2.1.2. Objectives	61
2.1.2.1. Resource Goals and Definitions	61
2.1.2.2. Specific Aims	64
2.1.2.3. Resource Scope	68
2.1.2.4. Significance to Biomedicine	69
2.1.3. Background	72
2.1.4. Resource Progress	74
2.1.4.1. Summary of Prior Goals	74
2.1.4.2. Progress Highlights	76
2.1.4.3. Resource Equipment Details	79
2.1.4.4. Core System Development	87
2.1.4.5. Core AI Research	99
2.1.4.6. Dissemination Activities	109
2.1.4.7. Training Activities	112
2.1.4.8. Resource Community Management	115
2.2. Planned Resource Activities	118
2.2.1. Core Research and Development	118
2.2.1.1. ONCOCIN-Related Core Research	120
2.2.1.2. Basic Research in AI	133
2.2.1.3. Resource Hardware and Core System Development	162
2.2.2. Collaborative Research	178
2.2.3. Service	178
2.2.4. Training	178
2.2.5. Dissemination	178
2.3. Resource Organizational Structure	179
2.3.1. Organizational Structure	179

Collaborative Project Abstracts

2.3.2. Resource Staff Responsibilities	180
2.3.3. Resource Operating Procedure	181
2.3.3.1. New Project Recruiting	181
2.3.3.2. Stanford Community Building	181
2.3.3.3. Existing Project Reviews	182
2.3.3.4. Resource Allocation Policies	182
2.3.4. Support of Service and Collaborative Projects	183
2.3.5. Resource Advisory Committee	184
3. Impact of Current Biomedical Problems	185
4. Institutional Development	187
5. Future Plans	189
6. Collaborative and User Projects	191
6.1. Stanford Projects	193
6.1.1. GUIDON/NEOMYCIN Project	194
6.1.2. MOLGEN Project	202
6.1.3. ONCOCIN Project	209
6.1.4. PROTEAN Project	220
6.1.5. RADIX Project	225
6.2. National AIM Projects	233
6.2.1. CADUCEUS Project	234
6.2.2. CLIPR - Hierarchical Models of Human Cognition	240
6.2.3. MENTOR Project	245
6.2.4. SOLVER Project	249
6.3. Stanford Pilot Projects	258
6.3.1. CAMDA Project	259
6.3.2. REFEREE Project	267
6.4. National AIM Pilot Projects	273
6.4.1. PATHFINDER Project	274
6.4.2. RXDX Project	279
Appendix A. Stanford Knowledge Systems Laboratory	285
Appendix B. Resource Operations and Usage Data	293
Appendix C. AIM Management Committee Membership	307
Appendix D. Collaborative Project Abstracts	311

List of Figures

Figure 1:	SUMEX-AIM DEC 2060 Configuration	83
Figure 2:	SUMEX-AIM DEC 2020 Configuration	84
Figure 3:	SUMEX-AIM Shared DEC VAX 11/780 Configuration	85
Figure 4:	SUMEX-AIM File Server Configuration	86
Figure 5:	SUMEX-AIM EtherNet Configuration	94
Figure 6:	Sample Chemotherapy Protocol Hierarchy	121
Figure 7:	Sample ONCOCIN Rule, Translated to English from Internal Format	122
Figure 8:	Components of the Meta-OPAL System	126
Figure 9:	Levels of BB1's Control Blackboard with Examples from PROTEAN	137
Figure 10:	An Example PROTEAN Heuristic at the Focus Level	138
Figure 11:	Example of Preliminary BB1 Explanation	139
Figure 12:	The components of a Learning System.	145
Figure 13:	Summary of Rules of Generalization and Specialization by Fu [19]	148
Figure 14:	Total CPU Time Consumed by Month	294
Figure 15:	Monthly CPU Usage by Community	296
Figure 16:	Monthly Terminal Connect Time by Community	297
Figure 17:	Cumulative CPU Usage Histogram by Project and Community	299
Figure 18:	TYMNET Terminal Connect Time	306
Figure 19:	ARPANET Terminal Connect Time	306