

## 자바컴포넌트 기반의 유전체 연구 협업지원 시스템

### (The Collaborative System for Genomic Research based on Java Component)

안건태\*, 김진홍\*, 한인섭\*\*, 이명준\*

\*울산대학교 컴퓨터정보통신공학부, \*\* 울산대학교 화학생명과학부

\*{java2u, avenue, mjlee}@mail.ulsan.ac.kr, \*\*hanis@mail.ulsan.ac.kr

#### 요 약

포스트지놈 시대가 도래되면서 유전체 관련 기술과 지원 프로그램에 대한 연구가 활발해지고 있다. 특히 유전체 연구의 특성상 연구그룹들이 지역적으로 분산된 경우가 많아서 연구과정에서 생성되는 결과물이나 정보들을 효과적으로 공유하고 교환할 수 있도록 지원하는 시스템의 필요성이 증대되고 있다. 이러한 시스템은 지속적인 서비스를 제공하면서도 사용자의 다양한 요구를 신속하게 처리하여 주는 것이 바람직하다.

본 논문에서는 유전체 연구를 위한 연구자에게 사용자 중심의 유전체 검색 시스템을 제공하고, 특정 연구 분야의 연구자 사이에 연구에 관련된 정보와 의사 교환을 용이하게 지원하는 BioPlace 협업지원 시스템에 대하여 기술한다. BioPlace 시스템은 자바 컴포넌트 아키텍처인 EJB 기술을 이용하여 확장성이 높으면서도 다중 사용자 환경에서 안정적인 서비스를 제공하도록 설계되었다.

구관련 정보와 의견을 효과적으로 공유하고 교환할 수 있도록 지원하는 시스템의 필요성이 증대되고 있다[1][2].

본 논문에서는 유전체 연구를 위한 연구자에게 유전체 검색 시스템을 제공하고, 특정 연구 분야의 연구자 사이에 연구에 관련된 정보와 의사 교환을 용이하게 지원하는 웹기반 공동작업 환경인 *BioPlace* 시스템에 대

#### 1. 서론

유전체 연구는 연구의 특성상 규모가 크고 지역적으로 널리 흩어져 있는 연구 기관들이 협력하여 연구를 수행하는 경우가 빈번히 발생한다. 따라서 특정 분야의 연구자들이 연

하여 기술한다. BioPlace 시스템은 자바 컴포넌트 기술을 기반으로 구현되었으며, 특정 분야에 관련된 유전체 연구자들이 사용자 중심의 인터페이스를 이용하여 효율적이고 안정적인 정보교환 및 유전체 연구 관련 서비스를 이용할 수 있는 기능을 지원한다. 또한, 기존에 개발된 많은 유전체관련 소프트웨어들이 리눅스나 유닉스 운영체제를 기반으로 하고 있어서 자바 프로그래밍언어로 개발할 경우 이러한 공개 프로그램들을 시스템에 용이하게 확장할 수 있는 장점이 있다. 따라서 본 논문에서는 개발이 용이하고 안정적이며 이식성이 강한 자바컴포넌트 아키텍처인 JSP와 EJB를 이용하여 BioPlace 시스템을 구현함으로써 다양한 유전체 연구 환경에 유연하게 대처할 수 있도록 하였다.

BioPlace 시스템에서는 연구자의 개인 데이터 관리를 위한 *Personal Workspace*와 관련 연구자들과 공동작업에 필요한 공동 데이터 관리를 위한 *Team Workspace*를 제공하고 있으며, 유전체 연구팀들이 *Team Workspace*에서 효율적으로 작업할 수 있도록 Research road map과 K-Blast 서비스를 제공한다. Research road map은 일반적인 유전체 연구에서 유용하게 사용될 수 있는 연구 지침서로서 연구에 필요한 자원이나 연구 단계에 대한 지시사항들을 팀작업장 내에 구현하여 추가하였다. K-Blast 서비스는 기존의 NCBI에서 제공하는BLAST 프로그램 [3][4]을 지역화(localization)하여 한글화된 사용자 인터페이스를 제공하여 국내 연구자들이 효율적으로 사용할 수 있도록 하고 있다.

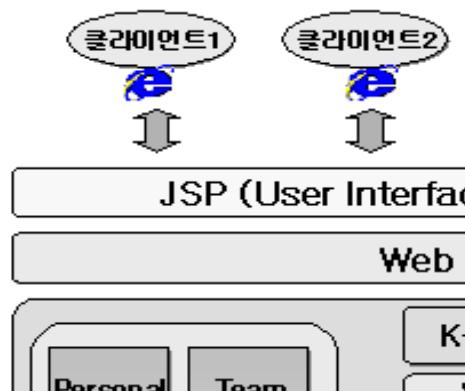
본 논문의 구성은 다음과 같다. 1장 서론에 이어 2장에서는 BioPlace 협업지원 시스템의 개요와 구성에 대하여 설명한다. 3장에

서는 BioPlace 시스템의 구현과 사용자 인터페이스에 대하여 기술하고 4장에서 결론을 맺고자 한다.

## 2. BioPlace 협업지원 시스템

생명체로부터 얻어낸 데이터로부터 실생활에 유용한 데이터를 얻기 위한 도구의 제작과 이를 이용한 유전체 연구는 다량의 생물학적 데이터를 분석할 수 있어야 하며, 생명정보, 생명공학, 생명화학 등 다양한 개별 연구 분야의 연구자들 사이에 정보를 교환하고 의사를 나눌 수 있는 도구에 대한 연구의 필요성이 증대되고 있다.

BioPlace는 이러한 목적을 위하여 설계된 시스템으로서, 유전체 연구를 위한 국내 연구자들에게 한글화된 사용자 중심의 유전체 검색엔진을 제공하고, 공동의 목적을 가진 연구자들 사이의 효과적인 정보의 공유와 교환을 지원하는 웹기반 협업 지원 시스템이다.

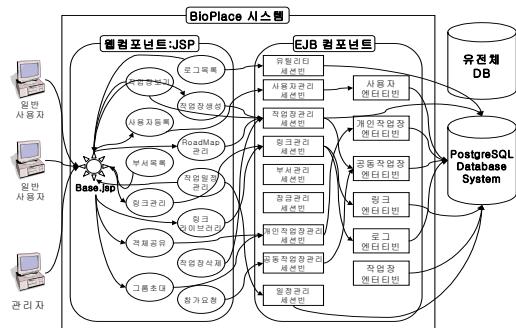


(그림 1) BioPlace 전체 시스템 구성

[그림 1]은 BioPlace의 전체 시스템 구성을 도식화한 것이다. 그림에서 보는 바와 같

이 BioPlace는 효과적인 협업 지원을 위하여 Workspace, Work Scheduler, K-Blast 검색 엔진 등을 하부 모듈로 제공한다. BioPlace 시스템의 사용자는 Web Browser를 통하여 시스템에 접근하여 원하는 작업을 수행할 수 있으며, 사용자들의 데이터 저장과 공유의 기능을 수행하는 Workspace는 대표적인 분산 아키텍처의 하나인 EJB (Enterprise Java Beans) 기술[5][6]을 이용하여 구현됨으로써, 다수의 사용자들에 대하여 안정적이며 높은 확정성을 지원하고 있다.

[그림 2]는 BioPlace 시스템을 구성하는 웹컴포넌트와 EJB 컴포넌트를 중심으로 시스템을 도식화하여 표현한 것이다. 추후 시스템에는 단백질 관련 데이터베이스와 이를 이용하는 컴포넌트들이 추가로 구현될 예정이다.



(그림 2) BioPlace 시스템의 기본 프레임워크

[그림 2]에서 보는 바와 같이 BioPlace 캠포넌트들은 기능상으로 분류가 되어 있어 시스템의 확장이나 변경사항이 발생시 유연하게 대처 가능한 구조로 설계되어 있다. 아래에서는 BioPlace를 구성하는 기본 내부 시스템들과 그들이 가진 기능과 역할에 대하여 기술한다.

## 2.1 WorkSpace

공동작업을 위한 연구자들이 효과적으로 연구자료를 공유하고 교환하기 위해서는 공동의 작업공간을 지원해 주어야 한다.

Workspace는 유전체 연구팀들의 원활한 협업을 지원하기 위한 공동작업 환경을 지원하며, 아울러 유전체 관련 작업의 성격에 따라 적절한 Research Road Map을 제시 하여 줌으로써 작업의 효율성을 증대시켰다. 추가로, Workspace는 사용자의 개인 정보관리를 목적으로 하는 Personal Workspace와 공동 작업에 필요한 정보를 공유하고 관리할 수 있게 설계된 Team Workspace 등의 두 가지 작업 공간을 지원한다.

## Personal Workspace

BioPlace 이용자는 인증과정을 통하여 BioPlace에 로그인하게 되면, 팀의 구성과는 관계없이 기본적으로 Personal Workspace를 할당 받게 된다. 각 사용자는 자신의 지역 컴퓨터로부터 Personal Workspace로 자료를 업로드하여 자신의 정보를 웹을 통하여 언제 어디서나 직접 관리할 수 있게 된다. Personal Workspace에 등록된 자료들은 차후 언제든지 지역 컴퓨터로 다운로드 될 수 있으며, 사용자 컴퓨터의 자료를 백업하는 용도로도 사용될 수 있다. Personal Workspace에서는 문서나 검색 결과물을 포함하여, 주요한 연구관련 링크, 메모 정보 등을 관리 가능하다.

## Team Workspace

Team Workspace는 공동의 과제를 수행하는 작업그룹을 위한 공간으로 공유 자료를 관리한다. 특정 Team Workspace에 참여하

고 있는 구성원은 자신의 Personal Workspace에서 관리되는 정보를 Team Workspace로 복사하거나 이동시킬 수도 있고, 직접 Team Workspace 내에서 새로 생성할 수도 있다. BioPlace 사용자는 공동과제의 수행을 위하여 작업그룹을 필요할 때마다 동적으로 생성할 수 있다.

작업그룹은 생성 시 선택적으로 다양한 성격을 지니도록 설정이 가능하다. 작업그룹을 생성한 사용자가 작업그룹의 관리자가 되며, 다른 사용자들을 작업그룹의 구성원으로 초청하는 과정을 통하여 그룹의 구성원을 조직한다.

작업그룹에 참여하는 방법은 그룹 관리자의 초청과 작업그룹에 참여하고자 하는 일반 사용자의 능동적인 그룹 참가 요청을 통하여 이루어진다. 또한, 구성된 팀의 작업그룹마다 토론을 할 수 있는 토론마당 기능과 공동작업 공간에서 발생하는 다양한 이벤트를 감시하기 위한 모니터링 기능도 제공한다.

[그림 3]은 Personal Workspace와 Team Workspace의 연관 관계를 다이어그램으로 표현한 것이다. 즉, 공동의 목적을 가진 작업그룹을 중심으로 그룹 구성원들이 유기적인 관계를 보인다.



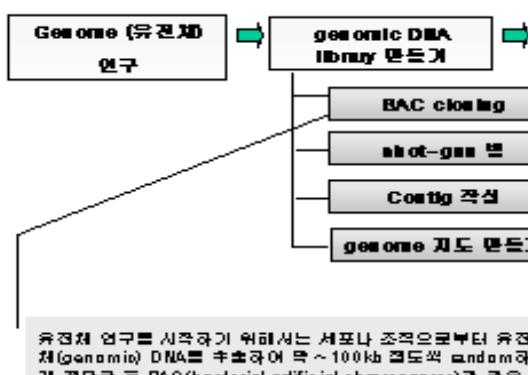
(그림 3) Workspace의 상호 연동

[그림 3]에서 보는 바와 같이 Workspace는 Personal Workspace와 Team Workspace의 공간이 유기적인 관계를 이루고 있으며, 작업장에 내에서 사용할 수 있는 다양한 객체들을 제공함으로써 보다 효율적인 정보의 공유와 관리기능을 지원한다. Workspace에서 제공하고 있는 객체 정보로는 폴더, 파일, 메모, URL 등이 있는데, 이러한 정보들은 필요에 따라 작업장 간의 이동이 가능하며, 또한, 개인 로컬 시스템으로 다운로드하거나 로컬시스템에서 Workspace로 원하는 자료를 업로드할 수 있는 기능을 가지도록 설계 및 구현되었다.

## 2.2 Research Road Map

유전체 연구는 다양한 생물학적 데이터를 각 연구 분야에 맞는 도구와 특정 분야의 자료를 바탕으로 총체적으로 이루어지고 있다.

현재 이러한 연구를 위한 도구들은 많이 제공되고 있으나 전문가를 제외한 일반 연구자들은 사용을 제대로 하지 못하는 실정이다. 이러한 문제를 해결하기 위하여 BioPlace에서는 유전체 연구에 있어서 특정 분야에 맞는 도구 및 유용한 자료를 쉽고 빠르게 찾아서 이용할 수 있도록 제작된 Research Road Map을 제공한다. 이 Research Road Map은 Workspace에 포함되어 있으며 특정 목적을 가진 작업그룹의 생성 시 BioPlace 사용자가 자신의 목적에 적합한 Research Road Map 구성을 할 수 있도록 되어 있다.



(그림 4) Genome Research Road Map

[그림 4]에서 보는 바와 같이 BioPlace 시스템에서는 유전체를 연구하는 그룹에서 연구팀의 연구방향과 절차를 제시할 수 있는 가장 일반적인 Genome Research Road Map을 제공하여 효과적인 연구 진행을 지원할 수 있도록 설계되었다.

### 2.3 Work Scheduler

Work Scheduler는 Workspace에서 이루어지는 다양한 일정을 각 구성원들에게 효과적으로 전달할 방법을 제공해 주는 시스템이다. 팀의 성격이나 일정의 성격에 맞게 다양한 일정분류체계를 지원하여 사용자들에게 보다 편리한 일정관리를 지원해 준다. 팀의 온라인 회의날짜, 정보교환 시기, 관련 분야의 실험 일정, 등 분류된 일정정보를 사용자의 일정표에 제공하고, 사용자의 요청에 따라 분류하여 보여주는 방식이다.

Work Scheduler는 특정 팀에 참여하고 있는 사용자의 리스트를 파악하여 각 사용자의 일정을 미리 알림으로써, 사용자는 Team Workspace에 접근하지 않고도 통합된 일정을 Personal Workspace에서 볼 수 있게 된

다.

일정관리 서비스를 통하여 각 사용자는 중복된 일정에 대한 조정을 보다 신속하게 처리하고, 각 일정에 대한 주석처리 또는 메모추가 기능을 통하여 세부사항을 정리할 수 있으며, 팀별 일정에 개인의 일정을 추가한 통합일정을 한눈에 볼 수 있어 능률적인 공동작업의 진행 계획을 세울 수 있다.

### 2.4 K-Blast

유전체 연구에서 데이터베이스 상동성 검색(database homology search) 위한 도구로 가장 많이 사용하는 것은 Blast이다. Blast는 DNA와 단백질 작은 조각의 정보를 가지고 이미 구축된 데이터베이스에서 그 조각이 기존의 정보 사이의 유전체 유사성검사나 유사한 유전자 사이에 또는 이 유전자가 코딩하는 단백질 사이의 구조관계 발견 등 유전체 연구의 많은 부분에서 필수적으로 사용되는 중요성이 상당히 높은 도구라 할 수 있다.

개발된 BioPlace에는 이러한 Blast 프로그램을 지역화(localization) 한 새로운 K-Blast 시스템을 탑재하였으며 한글화된 사용자 인터페이스의 구현과 Blast 도구의 사용법과 풍부한 사용예시를 제시함으로써 관련 사용자들의 편리를 도모하였다.

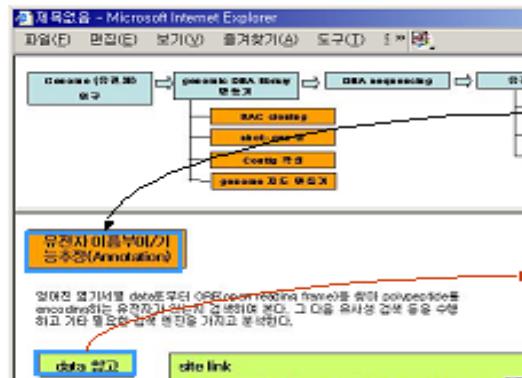
## 3. BioPlace 시스템의 구현

BioPlace는 유전체 관련 연구자들의 효과적인 정보교환 및 사용자 중심의 유전체 연구 관련 소프트웨어를 지원하는 통합 공동작업 환경이다. 시스템의 구현은 WorkSpace의 개발, Work Scheduler의 구현, 그리고 K-Blast 구축 등 세 단계로 나뉘어져 구현

되었으며, 최종적으로 각각의 모듈들을 BioPlace라는 이름으로 통합하였다.

### 3.1 WorkSpace 개발

본 논문에서는 웹을 통한 개인의 자료의 관리 및 공동작업을 위한 그룹의 생성이나 작업의 성격에 맞는 팀을 구성할 수 있도록 Workspace를 설계하고 구현하였다.



(그림 5) WorkSpace Interface

[그림 5]는 BioPlace 사용자가 생성한 WorkSpace의 인터페이스를 보여준다. WorkSpace 서버를 구축하는데 필요한 데이터베이스는 공개용 데이터베이스인 PostgreSQL[7]을 사용하였으며, 운영체제에 상관없이 동작할 수 있도록 자바기술(java technology)을 이용하여 구현하였다.

Road Map 프레임 부분에서 원하는 작업을 클릭하게 되면 해당 작업에 대한 설명과 라이브러리, 그리고 작업의 결과를 저장하고 관리할 수 있는 Personal Workspace와 Team Workspace로의 링크가 제공된다. 작업결과는 바로 Data 창고 역할을 하는 WorkSpace로 보내지며, 차후 재 참조하거나 자신의 로컬 컴퓨터로 다운 받을 수 있는 기능을 지원하고 있다.

### 3.2 Work Scheduler의 구현

본 논문에서는 개인의 일정관리 및 공동작업의 효과적인 수행을 지원하기 위하여 작업의 진행계획, 일정, 변경통보 등, 공동작업과정에서 발생할 수 있는 전달사항이나 계획사항을 통합해서 관리해주는 Work Scheduler를 구현하였다.



(그림 6) Work Scheduler 인터페이스

[그림 6]에서 보는 바와 같이 Work Scheduler에는 일정을 추가, 삭제, 수정, 그리고 검색할 수 있는 기능이 있으며, Personal Workspace나 Team Workspace에 있는 링크를 클릭하게 되면, 시스템은 [그림 6]과 같은 Work Scheduler 인터페이스 보여주게 되고 사용자는 원하는 일정관리 작업을 수행 할 수 있도록 하였다.

### 3.3 K-Blast 검색엔진

K-Blast 시스템은 기존 Blast도구를 한글화하여 사용자 중심의 인터페이스로 추가한 K-Blast 시스템을 구축하였다.

K-BLAST는 유전체 연구에서 가장 흔하

게 쓰는 서열비교분석 프로그램이다. NCBI에서 웹으로 제공하는 BLAST 사용자 인터페이스를 한글화하였으며, [그림 7]과 같은 초기화면을 가진다.



(그림 7) K-Blast 초기 메인 화면

원래 NCBI에서 제공되는 BLAST는 전문가들에게는 용이하나 익숙하지 않은 초보자들에게는 사용에 어려움을 겪게 된다. 이를 위하여 K-BLAST는 연구 초보자들도 쉽게 따라 갈 수 있는 '초보자용 따라하기' 기능도 포함하고 있다. 또한, [그림 8]에서 보는 바와 같이 결과의 이해를 돋기 위하여 결과읽기에 대한 내용을 추가로 제공하고 있어서 결과내용에 대한 보다 용이한 분석과 이해가 가능하리라 기대한다.

```

6. 그 아래에 각 data의 서열이
    염기서열 또는 아미노산 단위!
>gill19923482|ref|NM_016322.2|
family (RAB14), mRNA
Length = 4106+
+
Score = 200 bits (101), Expect
Identities = 344/425 (80%)
Strand=Plus / Plus+

```

검색 요구한 서열▶

Query: 98 gcaccatataacttatccc

(그림 8) 초보자용 따라하기 화면

끝으로, 개발된 서브 시스템들을 통합하여 웹기반 통합 작업 환경인 BioPlace 시스템을 구축하였다.

#### 4. 결론

본 논문에서는 유전체 연구의 효율적인 지원을 위한 웹기반 협업지원 시스템인 BioPlace의 개발에 대하여 기술하였다.

BioPlace는 유전체 연구를 위한 연구자들에게 국내 사용자 중심의 유전체 검색 시스템을 제공하고, 특정 연구 분야의 연구자 사이에 연구에 관련된 정보와 의사 교환을 용이하게 지원하기 위한 웹기반 공동작업 지원 시스템이다.

BioPlace는 자바컴포넌트 기술을 기반으로 개발됨으로써 기 개발된 CoWare[1,2]와 같은 윈도우 기술 중심의 시스템보다 구현이나 확장면에서 용이하며, 다른 소프트웨어들과의 유동적인 결합이 가능하여 관련 연구의 효율을 극대화 할 수 있도록 지원한다.

BioPlace 시스템에서는 유전체를 연구하는 그룹에서 연구팀의 연구방향과 절차를 제시할 수 있는 가장 일반적인 Research Road Map을 제공해 줌으로써 연구의 목표를 달성하는 시간을 단축하여 연구를 효율적으로 진행시킬 수 있도록 하였다. 또한, 유전체 연구의 많은 부분에서 필수적으로 사용되는 blast 프로그램을 지역화하고 한글 인터페이스화한 K-Blast 시스템을 BioPlace에서 지원하고 있다.

현재까지, 국내에서의 유전체 연구는 매우 숙련된 연구자만이 새로운 DNA 서열 및 단백질 기능을 알아내기 위하여 해당 연구에 관련된 외국 사이트에 접근하여 서비스를 이

용하는 형편이다. 따라서 사용자 중심의 한글화 인터페이스를 제공하고 효과적인 정보교환 및 사용자 중심의 유전체 연구 관련 소프트웨어[8]를 지원하는 웹기반 공동 작업환경인 BioPlace의 개발은 국내 유전체 연구의 활성 및 발전에 상당한 도움을 줄 것으로 전망된다.

향후 연구로는 단백질 구조 데이터베이스인 PDB에 대한 검색을 지원하는 K-PDB와 유전체 연구에 자주 참조되는 공개 프로그램들을 추가로 지원할 예정이다. 또한, P2P[9] 기반의 메신저시스템을 개발하여 팀원들 간의 실시간 메시지 교환이나 파일교환을 용이하게 하고, Research Road Map에 대한 풍부한 사용 예시를 제작함으로써 시스템의 활용성을 높일 예정이다.

## 감사의 글

본 연구는 정보통신부의 "정보통신기초기술연구지원사업"의 지원으로 수행되었습니다.

## 참고문헌

- [1] 문남두, 안건태, 김진홍, 한천용, 정명희, 이명준, "CoWare :효과적인 공동작업을 위한 웹기반 그룹웨어", 한국정보처리학회논문지 B 제 8-B권 제3호, pp.269~282, 2001.
- [2] Myung-Joon Lee, Chun-Yong Han, Geon-Tae Ahn, Jin-Hong Kim, Nam-Doo Moon and Myung-Hee Jung, "CoWare: A Web-Based Groupware for Effective Collaboration", Proceedings The 4th Korea-Russia International Symposium on Science and

Technology, Part3, pp.128~133, 2000.

[3] <http://www.ncbi.nlm.nih.gov/BLAST/>.

[4] Altschul S. F., Gish W., Miller W., Lipman D. J, Basic Local Alignment Search Tool, Journal of Molecular Biology, 1990

[5] Beca L., Fox G.C., Podgorny M., Component Architecture for Building Web-based Synchronous Collaboration Systems, Proceedings, IEEE 8th International Workshops on 1999.

[6] Java2 Platform, Enterprise Edition Blueprints, "J2EE™ Design Patterns".

[7] <http://www.postgresql.org>.

[8] <http://genomics.com/software/>.

[9] John G, P2P Application:New Internet Bandwidth Monsters, Garntner Group Research Note Tactical Guidelines, 2000.

## 안건태



1999년 2월 울산대학교 전자계산학과 졸업(학사)

2001년 2월 울산대학교 컴퓨터 · 정보통신공학부 졸업(석사)

2003년 2월 ~ 울산대학교

컴퓨터 · 정보통신공학부 공학박사과정 수료  
관심분야 : 생물정보학, 협업지원 시스템, 분산시스템, 이동에이전트 시스템 등.

### 김진홍



1999년 2월 울산대학교 전  
자계산학과 졸업(학사)  
2001년 2월 울산대학교 컴퓨터정보통신 공학부 졸업  
(석사)  
2003년 2월 ~ 울산대학교

1993년 8월 ~ 1994년 7월 미국 버지니아대

학 교환교수

관심분야 : 프로그래밍언어, 분산 객체 프로그래밍 시스템, 병행 실시간 컴퓨팅, 인터넷 프로그래밍시스템, 생물정보학 등.

컴퓨터정보통신 공학부 박사과정 수료  
관심분야 : 생물정보학, 제한프로그래밍, 협  
업지원 시스템, 이동에이전트 시스템 등.

### 한인섭



1983년 서울대학교 이학사  
1988년 조오지아대학교 이학  
석사  
1988년 조오지아대학교 이학  
박사  
1988년 9월 ~ 1991년 1월

Univ. of California, Irvine (post-doc)

1991년 3월 ~ 현재 울산대학교 교수

관심분야 : 생물정보학, 면역학, 유전체학

### 이명준



1980년 2월 서울대학교 수  
학과 졸업(학사)  
1982년 2월 한국과학기술원  
전산학과 졸업(석사)  
1991년 8월 한국과학기술원  
전산학과 졸업(박사)

1982년 3월 ~ 현재 울산대학교 컴퓨터정보  
통신 공학부(교수)

