



Illinois Genetic Algorithms Laboratory  
Department of General Engineering  
University of Illinois at Urbana-Champaign  
Urbana, IL 61801.

# **DISCUS: Distributed Innovation and Scalable Collaboration in Uncertain Settings**

**David E. Goldberg,**

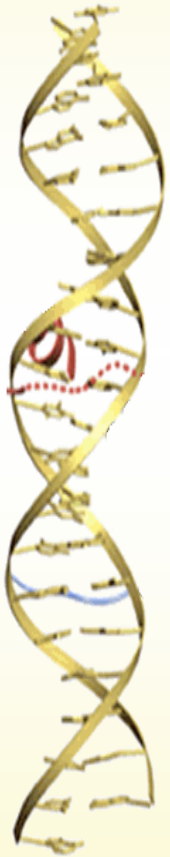
**Michael Welge, & Xavier Llorà**

NCSA/ALG + IlliGAL

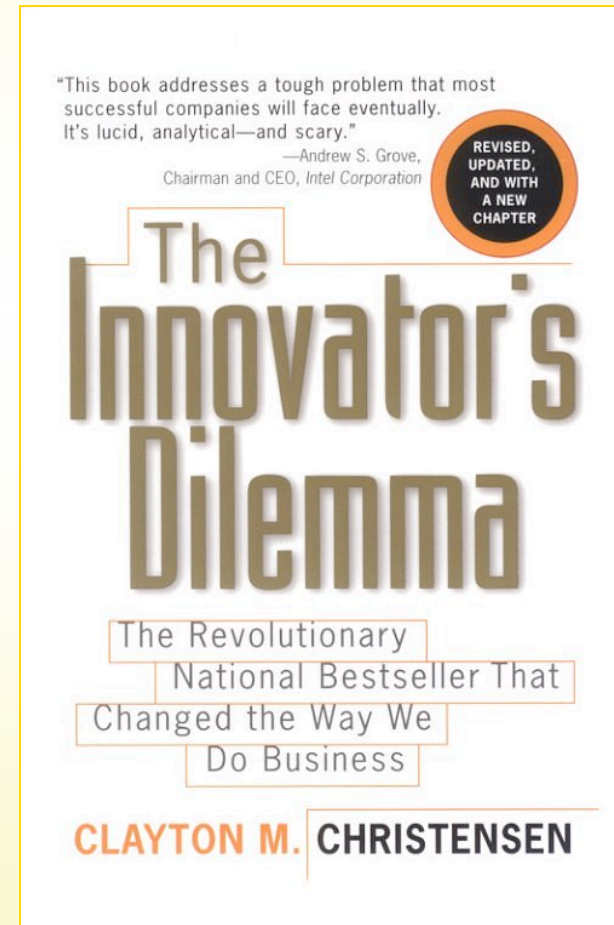
University of Illinois at Urbana-Champaign

`deg@uiuc.edu`

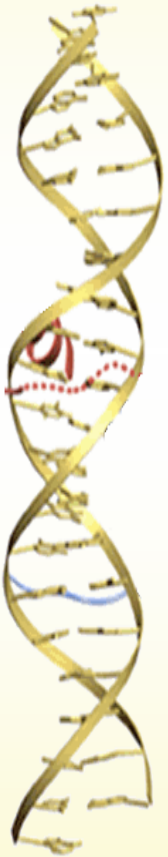
# Innovation This & Innovation That



- The business world is abuzz with “innovation.”
- Popular books tell companies how to get it.
- But little scientific understanding of what it is.
- UIUC research changing that.



# From Decision Support & Knowledge Management to Innovation Support



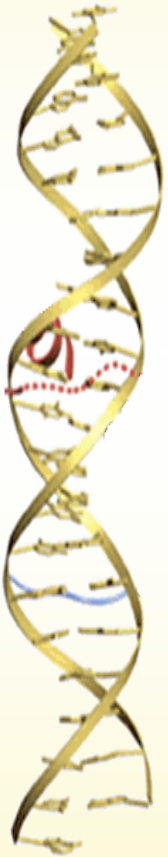
- Decision support systems help evaluate enumerated alternatives.
- Knowledge management helps manage that which is known.
- Can we build on DSS & KM to create *innovation support system* to systematically permit organizations to use IT to support pervasive and persistent innovation to their competitive advantage?

# Collaboration + Key Ideas = Opportunity



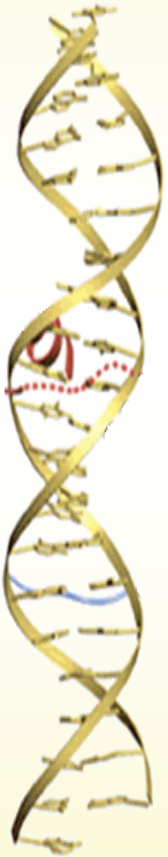
- Previous collaboration of ALG + IlliGAL
  - Applications-ready GA theory
  - MOGAs for D2K & the real world
  - Interactive genetic algorithms
- Confluence of key ideas
  - Interactive GAs
  - Human-based GA (Kosorukoff & Goldberg, 2002)
  - Chance discovery & data-text mining
- **DISCUS: Distributed Innovation and Scalable Collaboration in Uncertain Settings**

# Overview



- 3 elements research from IlliGAL
- 4 trips to the South Farms
- 2 trips to Japan
- The innovation connection
- The key problem: interactive superficiality
- KeyGraphs as aid to reflection
- Key elements of DISCUS
- Progress to date and anticipated

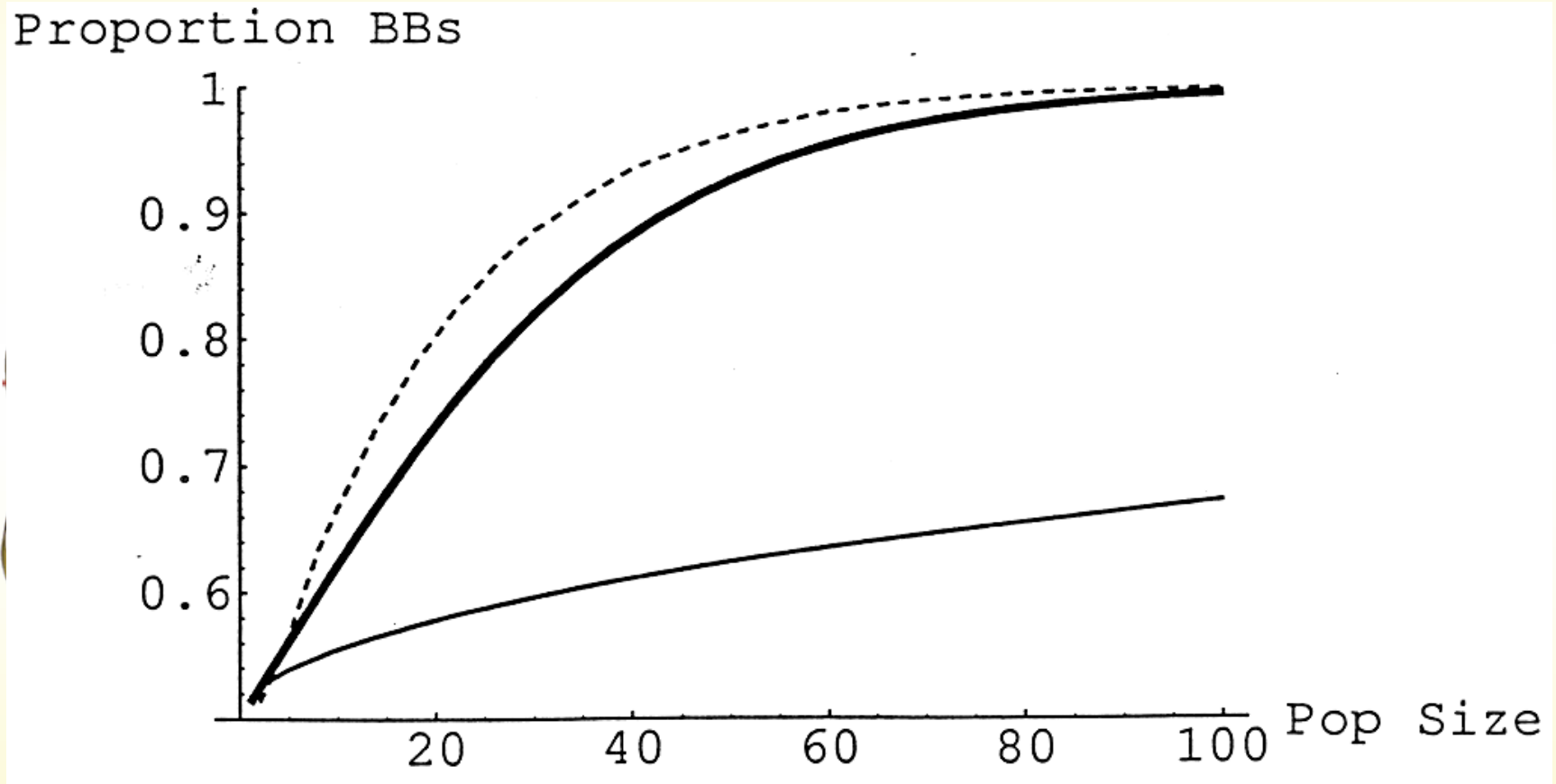
# 3 Elements from IlliGAL



- IlliGAL has studied principled
  - Genetic algorithm design theory
  - Genetic algorithm competence
  - Genetic algorithm efficiency
- Design theory permits analysis w/o tears.
- Competence = solve hard problems, quickly, reliably, and accurately  $\rightarrow O(l^2)$ .
- Efficiency takes *tractable* (subquadratic) solutions to *practicality*.

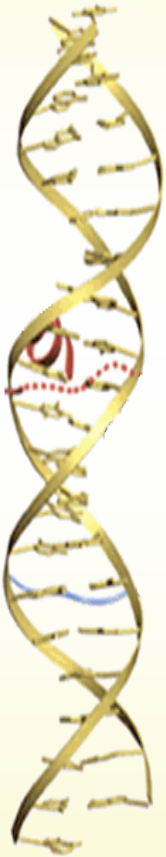


# GA Design Theory Makes Time and Quality Predictable

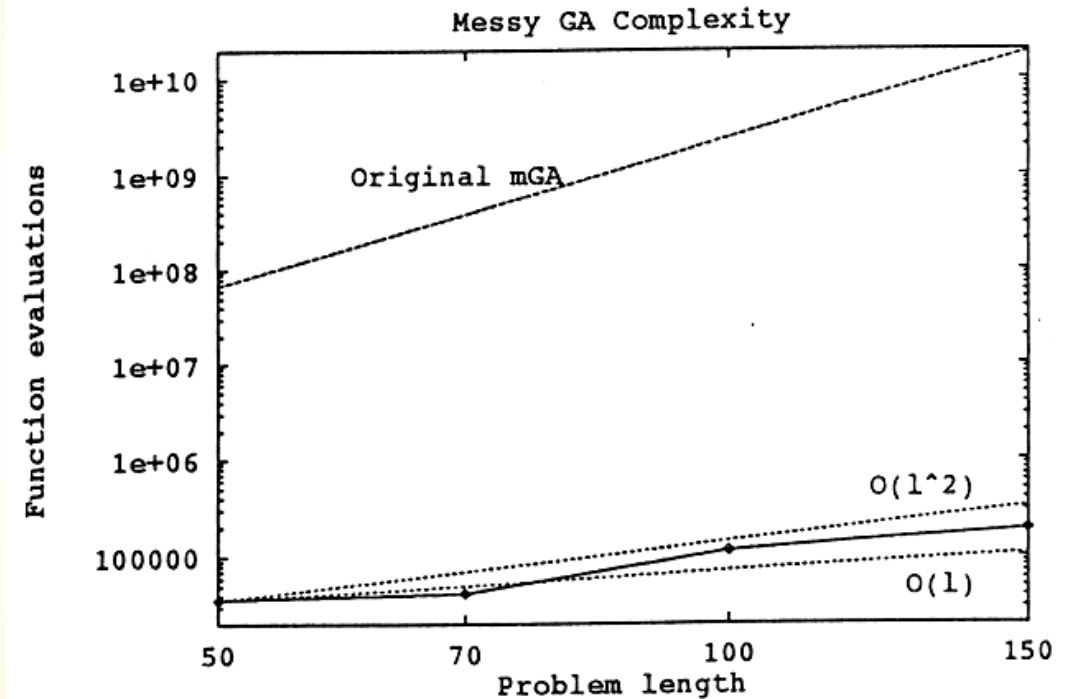




# 1993 Principled Scalable Computational Innovation Achieved



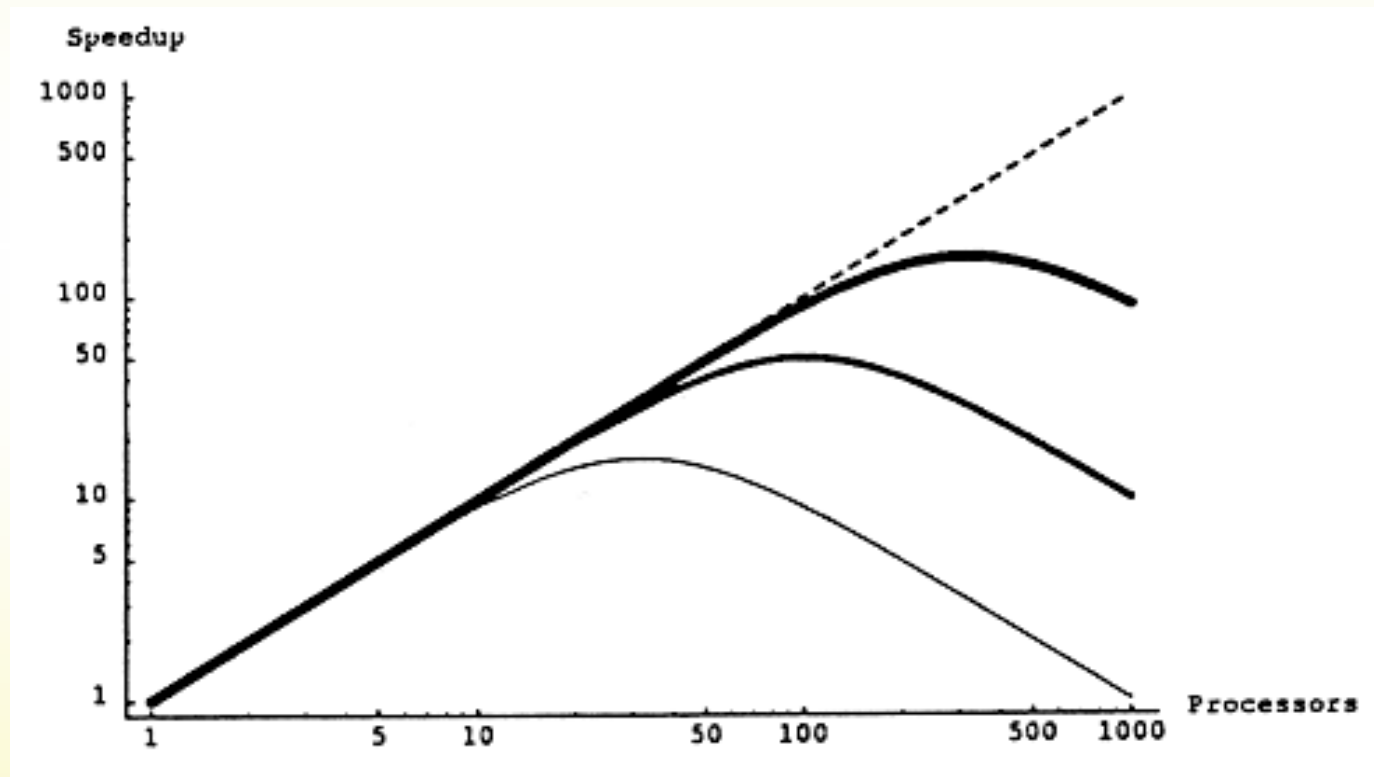
- Fast messy GA (1993) demonstrates principled, scalable innovation on hard problems.
- Subquadratic solutions
- 2001 - hBOA, hierarchical Bayesian optimization algorithm





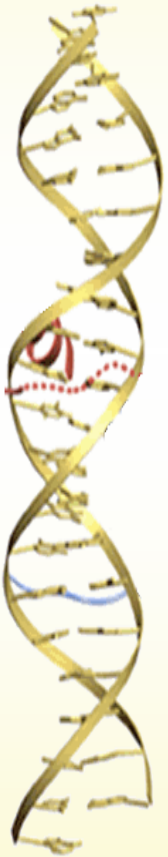
# Speedups and Efficiency

$$Sp = \frac{T_s}{T_p} = \frac{nT_f}{\frac{nT_f}{P} + PT_c} \quad \frac{T_f}{T_c} = 1, 10, 100$$



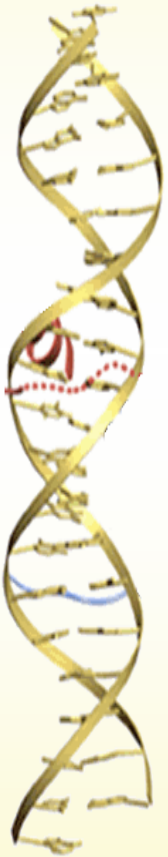
Optimal speedup  $0.5P^*$

# 4 Trips to NCSA South Farms



- Collaboration had blossomed with ALG & Prof. Minsker on
  - Carrying principled design theory to practice
  - Multiobjective selection to D2K & practice
  - GBML and HBGAs to D2K
  - Interactive GAs
- Keys for the current project:
  - HBGAs
  - Interactive GAs

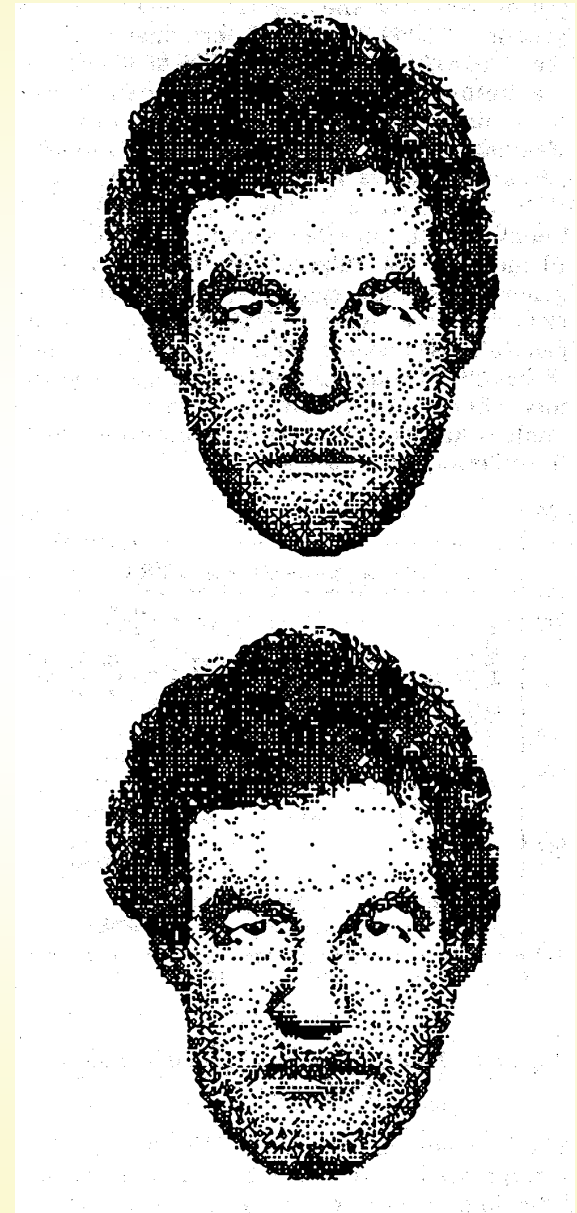
# Interactive & Human-Based Genetic Algorithms



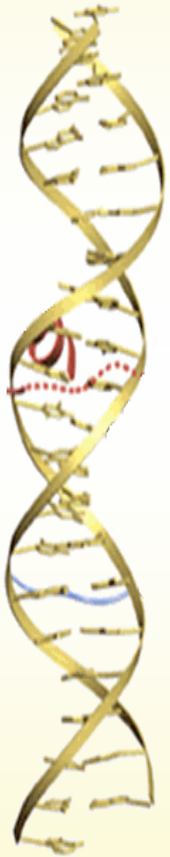
- Interactive GAs replace machine eval with human eval
- Human-based GAs replace ops & eval with human:

[www.3form.com](http://www.3form.com)

**Figure :** Actual photo of simulated criminal (above). Evolved image from witness using Faceprints (below).



# A Taxonomy of Evolutionary Methods Depending on Who/What Selects and Recombines

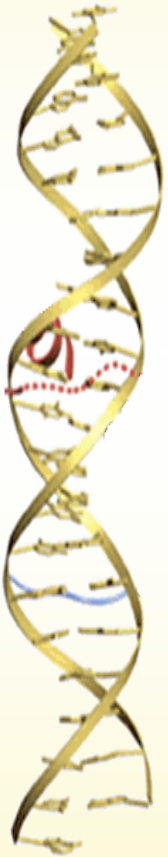


**Recombination agent**

<i>computational</i>	<b>Standard Genetic Algorithms</b>	<b>Interactive Genetic Algorithms</b>
<i>human</i>	<b>Computer Aided Design (CAD)</b>	<b>Human Based Genetic Algorithms</b>
	<i>computational</i>	<i>human</i>

**Selection agent**

## 2 Trips to Japan

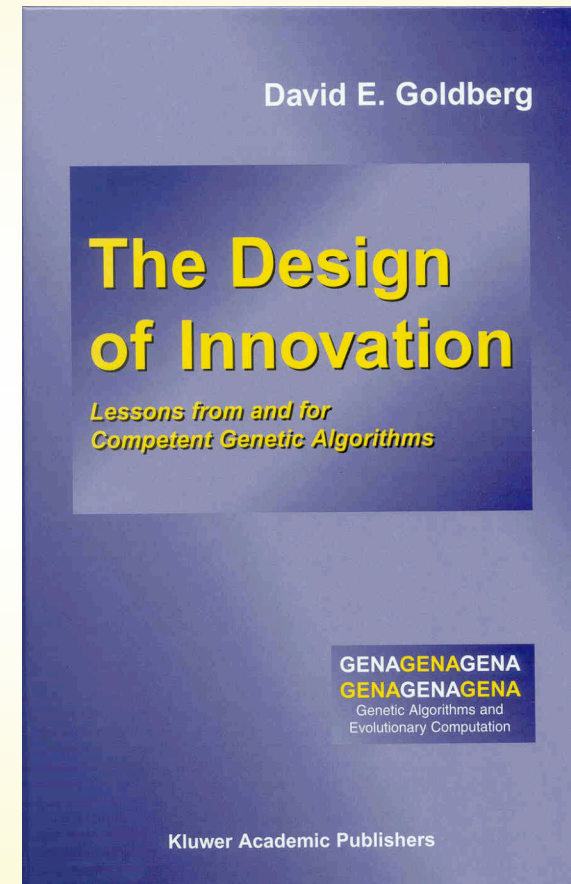


- Visited Tsukuba University, Graduate School of Systems Management, December 2001 – January 2002.
- Met KeyGraph Inventor & Chance Discovery Proponent, Yukio Osawa.
- Did Tutorial with Dr. Osawa August 2002.
- Finally understood importance of topic & relation to GAs.

# Modes of Innovation

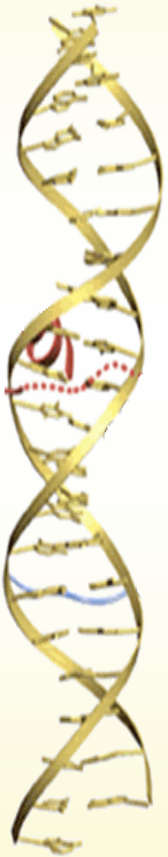


- GA as model of innovation
  - Kaizen = selection + mutation
  - Discontinuous change = selection + crossover
- Chance discovery
  - Low probability events linked to matters of importance
- Keygraphs as one computational embodiment of chance discovery.



<http://www-doi.ge.uiuc.edu>

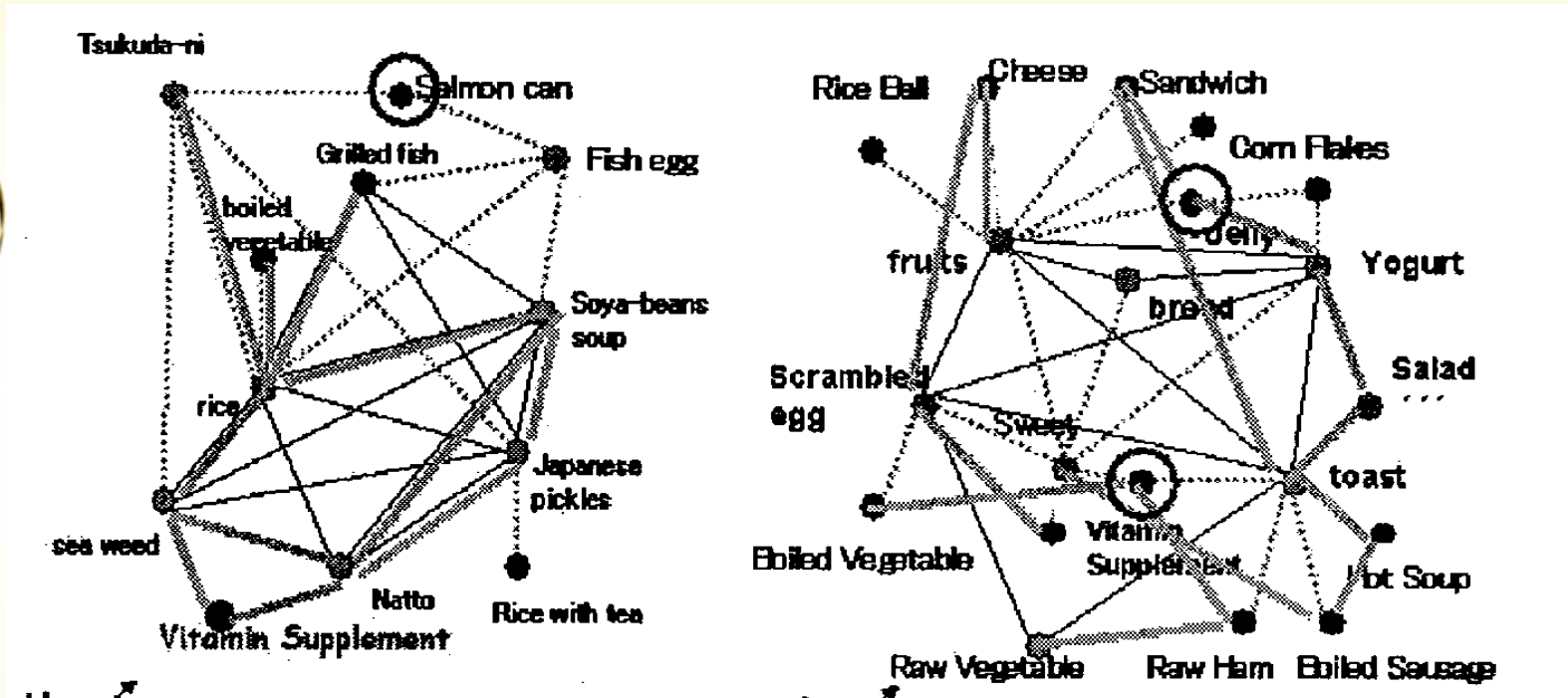
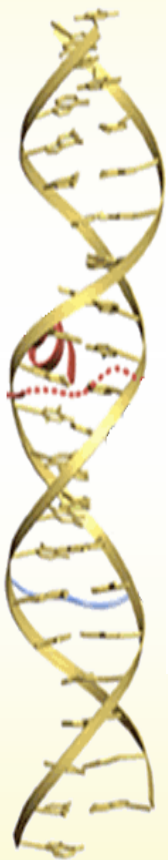
## Selection+Recombination = Innovation



- Combine *notions* to form *ideas* (Goldberg, 1983).
- “It takes two to invent anything. The one makes up combinations; the other chooses, recognizes what he wishes and what is important to him in the mass of the things which the former has imparted to him.”

*P. Valéry*

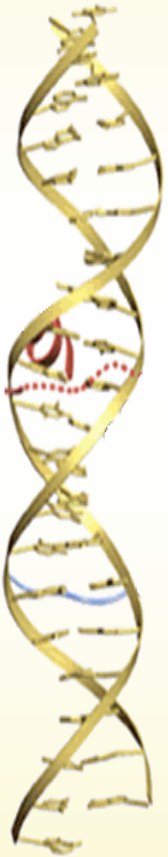
# KeyGraph Example: Japanese Breakfast



**Figure:** KeyGraph (Ohsawa, 2002) shows two clusters of food preferences for Japanese breakfast eaters. The chance discovery of rare use of vitamins was viewed as a marketing opportunity by food companies.

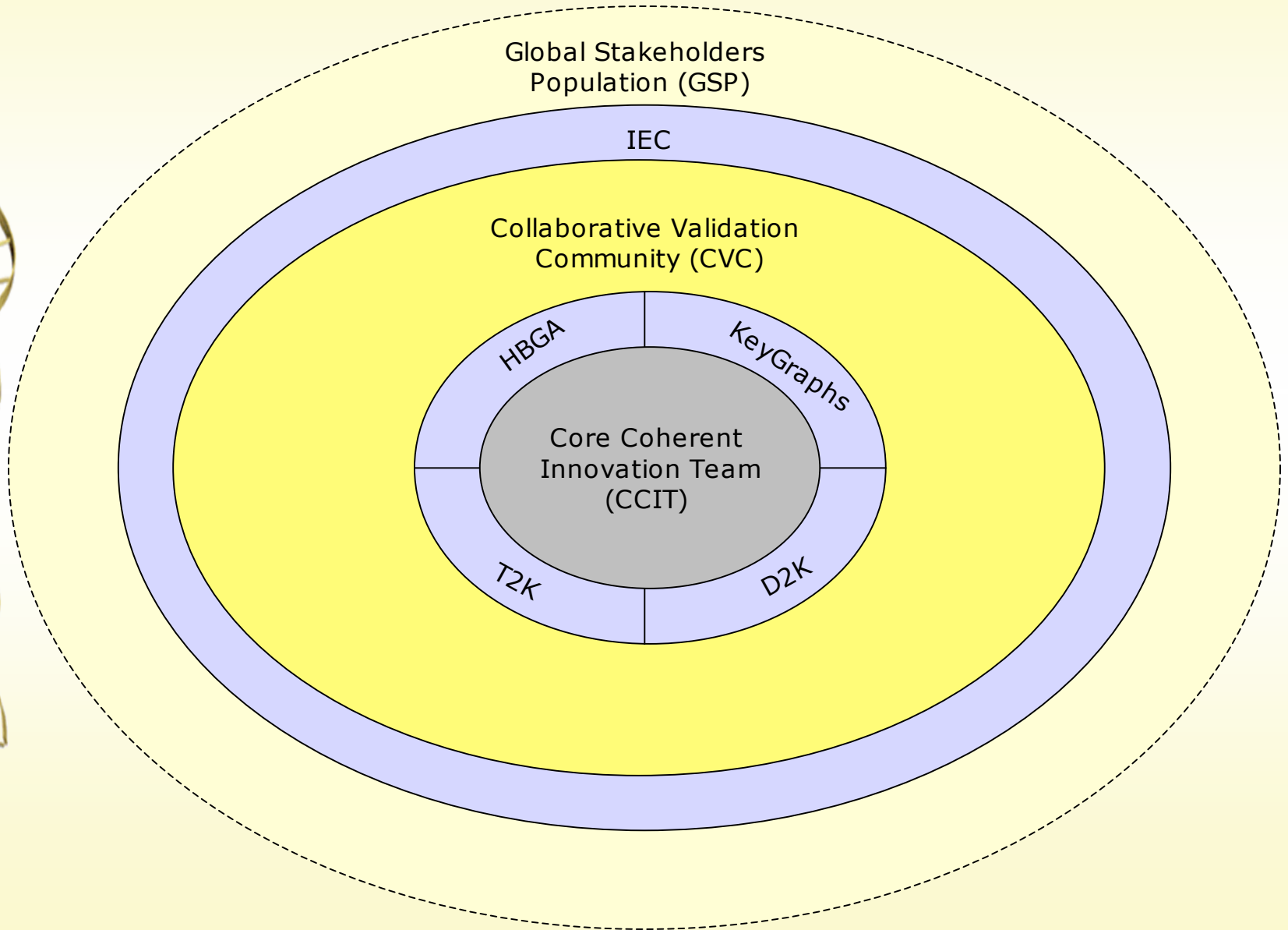


# Key Problem & Notion

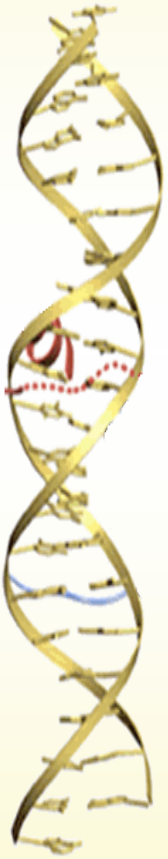


- Human-based GAs interesting, but suffer from *interactive superficiality*.
- KeyGraphs have been used to gain insight into text data, but usually *batch mode of processing*.
- **Combine interactivity of HBGAs and insight & reflection promotion of KeyGraphs.**
- Boost everything with competent efficient GAs and IEC at population outskirts.

# DISCUS Overall Design



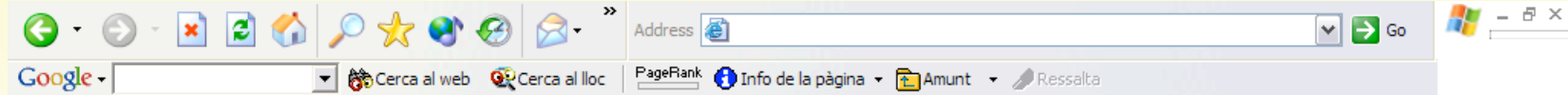
## Progress Now and Expected



- Started project from dead start in January 2003 (Dr. Xavier Llorca, Project Leader).
- Today: Have message board/chat/video conference + keygraph + rudimentary HBGA.
- June 2003: Start tests on internal problems solving.
- September 2003: Integrated pilot system.
- 2004: Looking for marketing & security applications.



Illinois Genetic Algorithms Laboratory  
Department of General Engineering  
University of Illinois at Urbana-Champaign  
Urbana, IL 61801.



# D.i.s.c.u.s.

[Home](#) | [Chat](#) | [Message Boards](#)

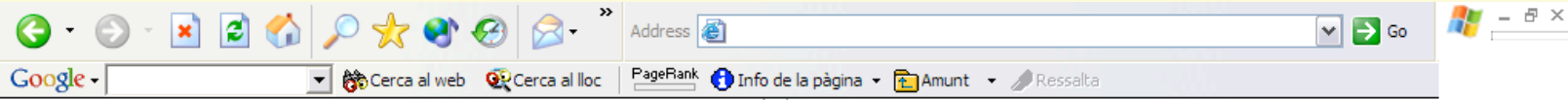
[Change Password](#) | [Logout](#)

Message boards:

- [Discus](#)
- [GA](#)
- [Experiment](#)
- [Genetic Algorithms and Programming: General Discussion](#)
- [Discovering Knowledge in BP Service Station Data](#)



Illinois Genetic Algorithms Laboratory  
Department of General Engineering  
University of Illinois at Urbana-Champaign  
Urbana, IL 61801.



# D.i.s.c.u.s.

[Home](#) | [Chat](#) | [Message Boards](#) [Logout](#)

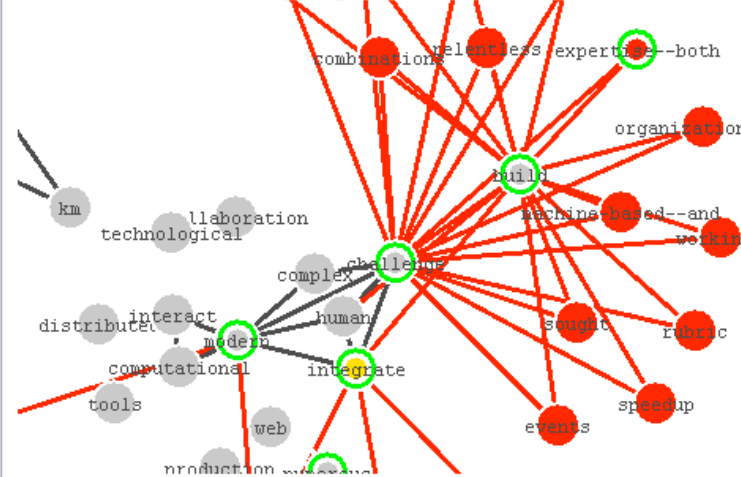
## Message Board > View Topic

# [Back To Forum](#) Topic: **Welcome** [Post A New Subject](#)

From: [Xavier](#) | Date: 2003-03-03 00:14:58.0 | Subject: [Welcome](#) | [View Replies](#) | [Reply](#)

Modern times challenge organizations and their leaders to adapt quickly and well to complex, fast-moving circumstances under trying conditions. Data sources are numerous, distributed, and contradictory. Challenges are difficult to detect and diagnose, widely dispersed, and constantly changing. Sources of knowledge and expertise are distributed, of varying quality, and difficult to integrate. Moreover, the tools of the trade are increasing in technological sophistication, computational intensity, and require specialized hardware, software, and care and feeding. Against this backdrop, modern technology has developed large-scale computer networks--most notably, the Internet--and the web has been developed at a rapid pace, allowing organizations to interact through web portals, e-mail, instant messaging, and other tools. These tools have had immediate impact in allowing individuals to communicate with one another conveniently and well. This has enabled the traditional means of human-to-human organizational collaboration to be carried out at a distance more effectively. Nonetheless, the sheer amounts of data available, the numerous sources of expertise--both human- and machine-based--and the relentless speedup of events threatens to challenge even these technological improvements to the workings of modern organizations. As a result, many have sought to build combinations of information technology under the rubric of knowledge management to support collaboration and the integration of multiple data sources. Loosely defined, knowledge management integrates IT and people to improve organizational learning and improvement. KM initiatives may be as simple as building databases

*(How to understand a use a Key Graph in a nutshell)*



*(KeyGrab zoom)*

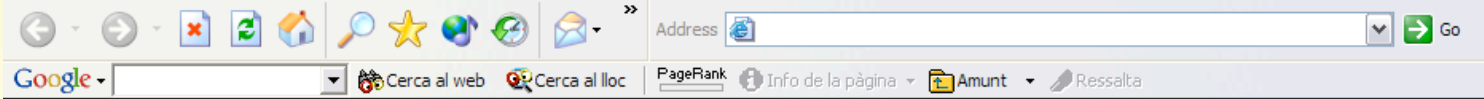
### Solution Center

- [Browse the solutions proposed so far for this topic](#)
- [Browse the proposed solutions by a particular user](#)
- [Rank of voted solutions](#)
- [Multivoting session for the current solutions](#)
- [Create a new free associative solution](#)
- [Create a new innovation-guided solution](#)

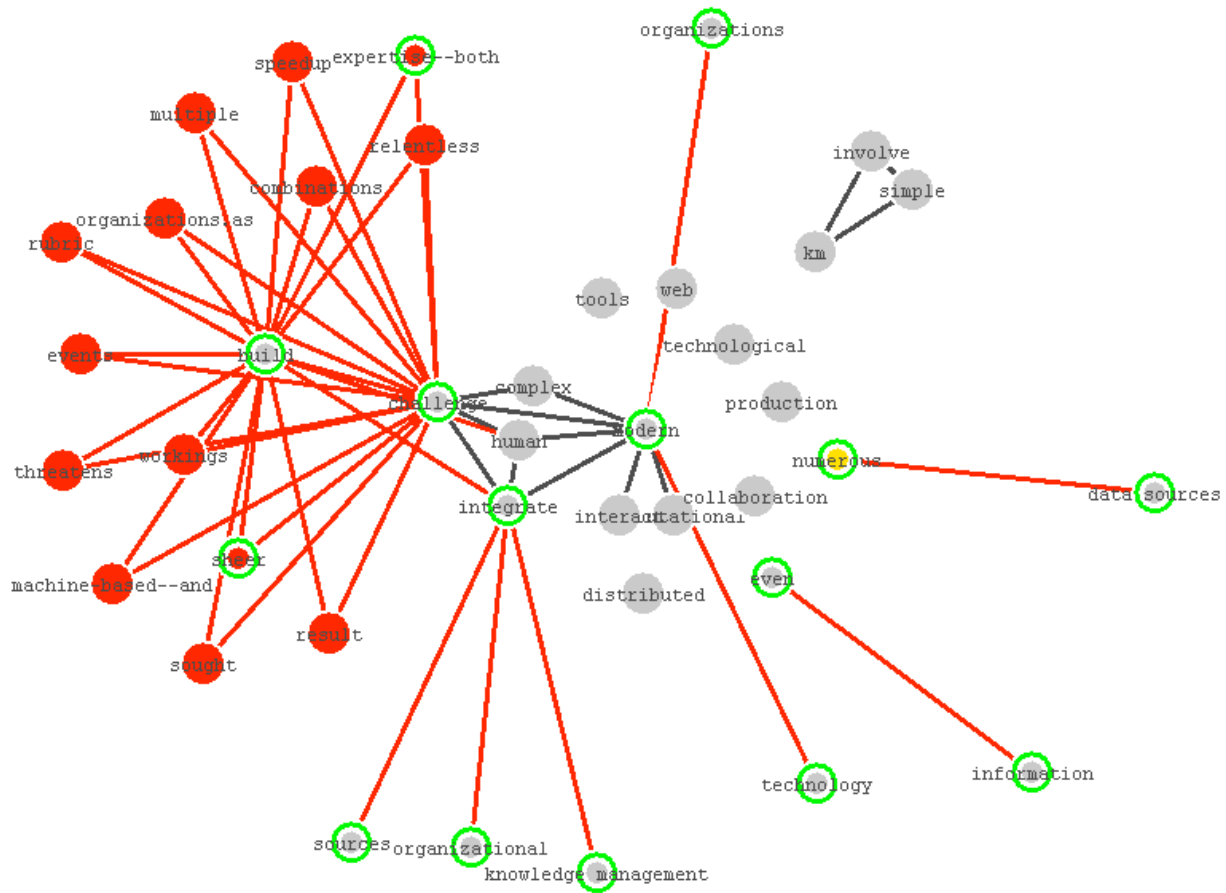
*(User: xllora) Tue May 20 09:49:51 CDT 2003*



Illinois Genetic Algorithms Laboratory  
Department of General Engineering  
University of Illinois at Urbana-Champaign  
Urbana, IL 61801.

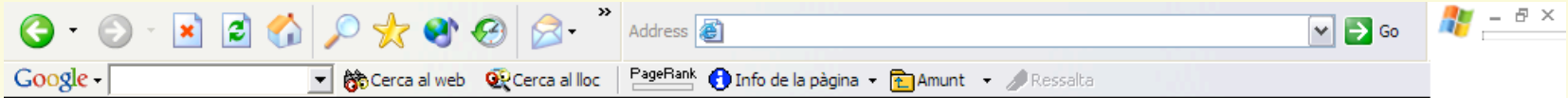


*(How to understand a use a Key Graph in a nutshell)*





Illinois Genetic Algorithms Laboratory  
Department of General Engineering  
University of Illinois at Urbana-Champaign  
Urbana, IL 61801.



D.i

Home | Cl

M

# Back 7

Forum

From:

Replie

Moder

and w

Data s

are dif

changi

quality

Showing the context of the innovation - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media

Address

Google Cerca al web Cerca al lloc PageRank Info de la pàgina Amunt Ressalta

### Solutions related to innovation

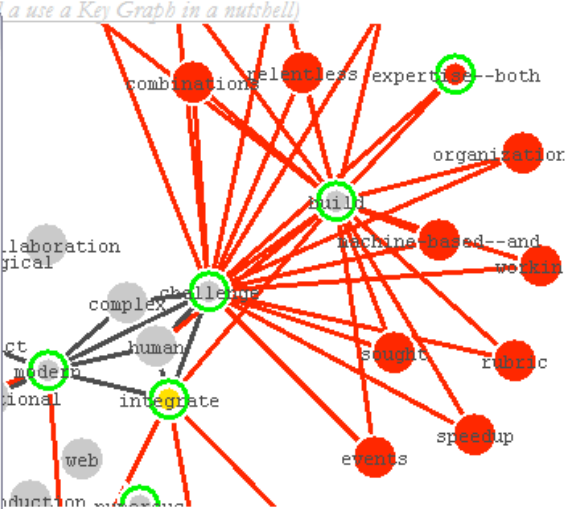
Solution 2 proposed by [xlora](#) on 2003-05-09 16:29:44.0

*New solutions will be created using human **innovation***

Solution 3 proposed by [xlora](#) on 2003-05-09 16:34:17.0

*Even small information data sources are relevant to **innovation***

Tue May 20 09:59:30 CDT 2003



increasing in technological sophistication, computational intensity, and require specialized hardware, software, and care and feeding. Against this backdrop, modern technology has developed large-scale computer networks--most notably, the Internet--and the web has been developed at a rapid pace, allowing organizations to interact through web portals, e-mail, instant messaging, and other tools. These tools have had immediate impact in allowing individuals to communicate with one another conveniently and well. This has enabled the traditional means of human-to-human organizational collaboration to be carried out at a distance more effectively. Nonetheless, the sheer amounts of data available, the numerous sources of expertise--both human- and machine-based--and the relentless speedup of events threatens to challenge even these technological improvements to the workings of modern organizations. As a result, many have sought to build combinations of information technology under the rubric of knowledge management to support collaboration and the integration of multiple data sources. Loosely defined, knowledge management integrates IT and people to improve organizational learning and improvement. KM initiatives may be as simple as building databases

### Solution Center

[Browse the solutions proposed so far for this topic](#)

[Browse the proposed solutions by a particular user](#)

[Rank of voted solutions](#)

[Multivoting session for the current solutions](#)

[Create a new free associative solution](#)

[Create a new innovation-guided solution](#)

(User: [xlora](#)) Tue May 20 09:49:51 CDT 2003



Illinois Genetic Algorithms Laboratory  
Department of General Engineering  
University of Illinois at Urbana-Champaign  
Urbana, IL 61801.

The screenshot shows a web browser window with two tabs. The active tab is titled "Dendrogram View - Microsoft Internet Explorer". The main content area displays a dendrogram with 25 clusters of terms. The terms are listed on the left, and the dendrogram branches to the right, with colors ranging from orange to green. Below the dendrogram are three radio buttons: "Cohesion" (selected), "Size", and "Size as Color". A "Close" button is at the bottom right of the dendrogram window.

The second tab is titled "Linkage Learning GA - Bibliography - Microsoft Internet Explorer". It shows a search result for "Linkage Learning Genetic Algorithms". The search results include:

- Text mining analysis of the... (requires a Windows browser)
- "Linkage Learning", Harik, ...
- "Learning Gene Linkage to Genetic Algorithms", Gold...

Summary statistics for the search results:

- 11 Groups
- 25 Clusters
- 196 Documents

A "bottom" button is visible below the statistics.

Below the search results, the title of the selected document is visible: "Compressed Introns in Learning Genetic Algorithms". The main text of the document is displayed below the title:

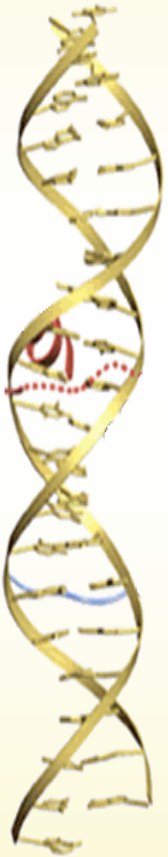
This paper revisited the linkage-learning genetic algorithm and showed that its time complexity is almost linear for exponentially scaled problems. Aside from that, this work represents a small step towards a competent GA. The efficient representation of the non-coding material is likely to be crucial for the solution of uniformly scaled problems within the linkage learning GA framework. Investigations in this direction have already begun and will be reported soon.

At the bottom right of the browser window, a list of documents is shown:

	doc147
701	<a href="#">Compressed Introns in a Linkage Learning Genetic Algorithm: Test problems</a>
	doc150
685	<a href="#">Compressed Introns in a Linkage Learning Genetic Algorithm: The limitations of SGAs</a>
	doc149

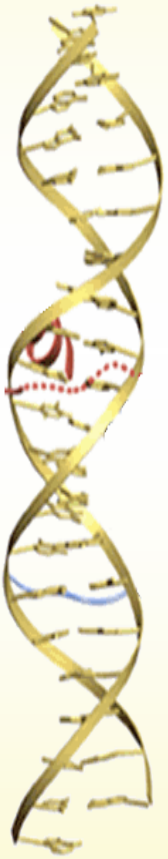


# Summary



- 3 imports from IlliGAL
- 4 trips to South Farms
- 2 trips to Japan
- The innovation connection
- Key problem: interactive superficiality
- Possible solution: interactive collaboration with reflection boosted by KeyGraphs
- Larger framework with competent & interactive GEC.

## Conclusions



- System emerging for innovation support.
- Envision both synchronous brainstorming and asynchronous continuing innovation.
- Combine HHC (human-human collaboration) and HMC (human-machine collaboration) to form powerful system.
- Overcome superficiality of online interaction through augmented reflection.
- Tackle challenge of the outer ring.

# Cast of 1000 Characters

**Powered by TRECC:** This project is supported by TRECC, a program of the UIUC administered by the NCSA by the Office of Naval Research under Grant #: N00014-01-1-0175

**IlligAL:** David E. Goldberg, Xavier Llorà, Kei Ohnishi, Tian Li Yu, Martin Butz, Antonio Gonzales

**ALG:** Michael Welge, Loretta Auvil, Duane Sears Smith, Bei Yu

**Knowledge and Learning System Group:** Tim Wentling, Andrew Wadsworth, Luigi Marini, Raj Barnerjee

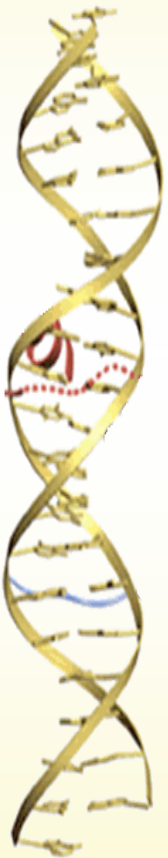
**Data Mining and Visualization Division:** Alan Craig

**Minsker Research Group:** Barbara Minsker, Abhishek Singh, Meghna Babbar

**Takagi Labo:** Hideyuki Takagi

**University of Tsukuba GSM:** Yukio Ohsawa

**Others:** Ali Yassine (GE), Miao Zhuang (GE), VIAS (the Visualization Information Archival/Retrieval Service)



## More Information



- Contact {xllora,deg}@illigal.ge.uiuc.edu
- Visit IlliGAL web site.
- <http://www-illigal.ge.uiuc.edu/>
- <http://www-discus.ge.uiuc.edu/>
- Recent book: Goldberg, D. E. (2002). *The Design of Innovation*. Boston, MA: Kluwer Academic, <http://www-doi.ge.uiuc.edu/>