The impact of internet technology on research and education in statistics

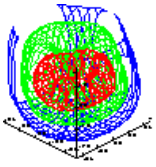
http://www.xplore-stat.de

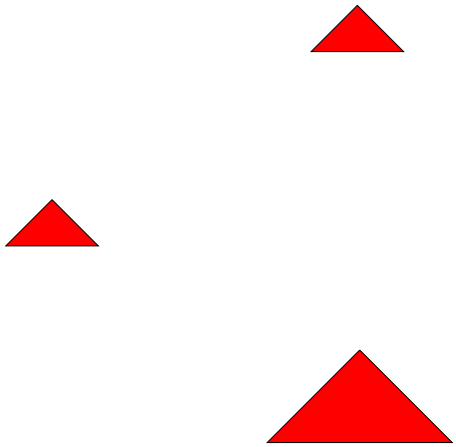
W. Härdle T. Kleinow

S. Klinke S. Schmelzer

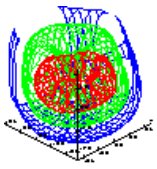
M. Müller

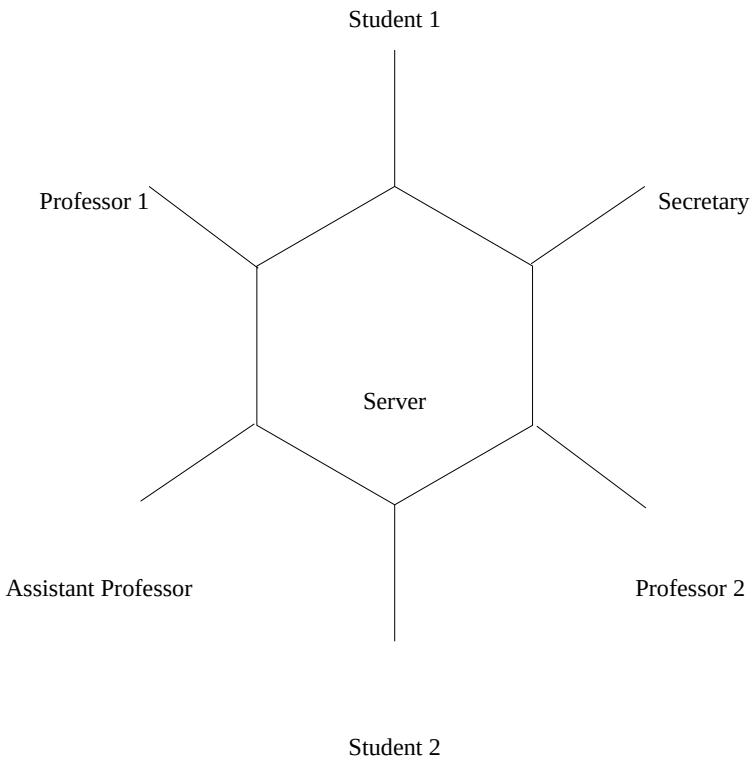
MD*Tech

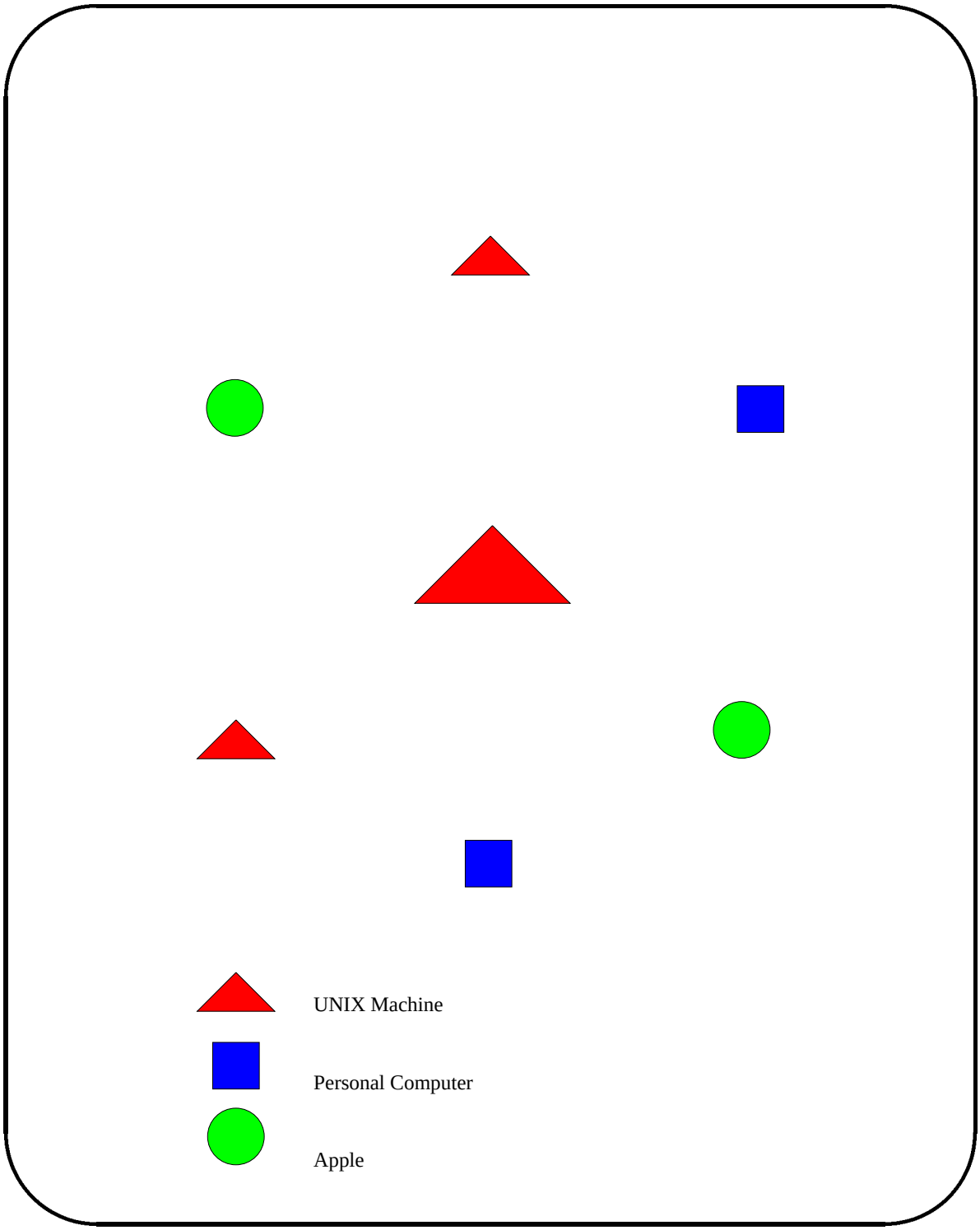


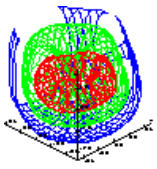


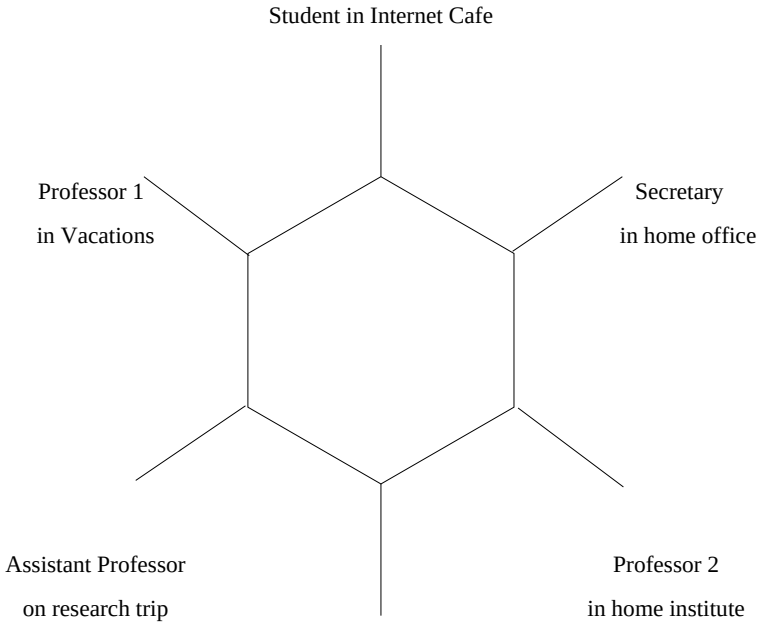
UNIX Machine

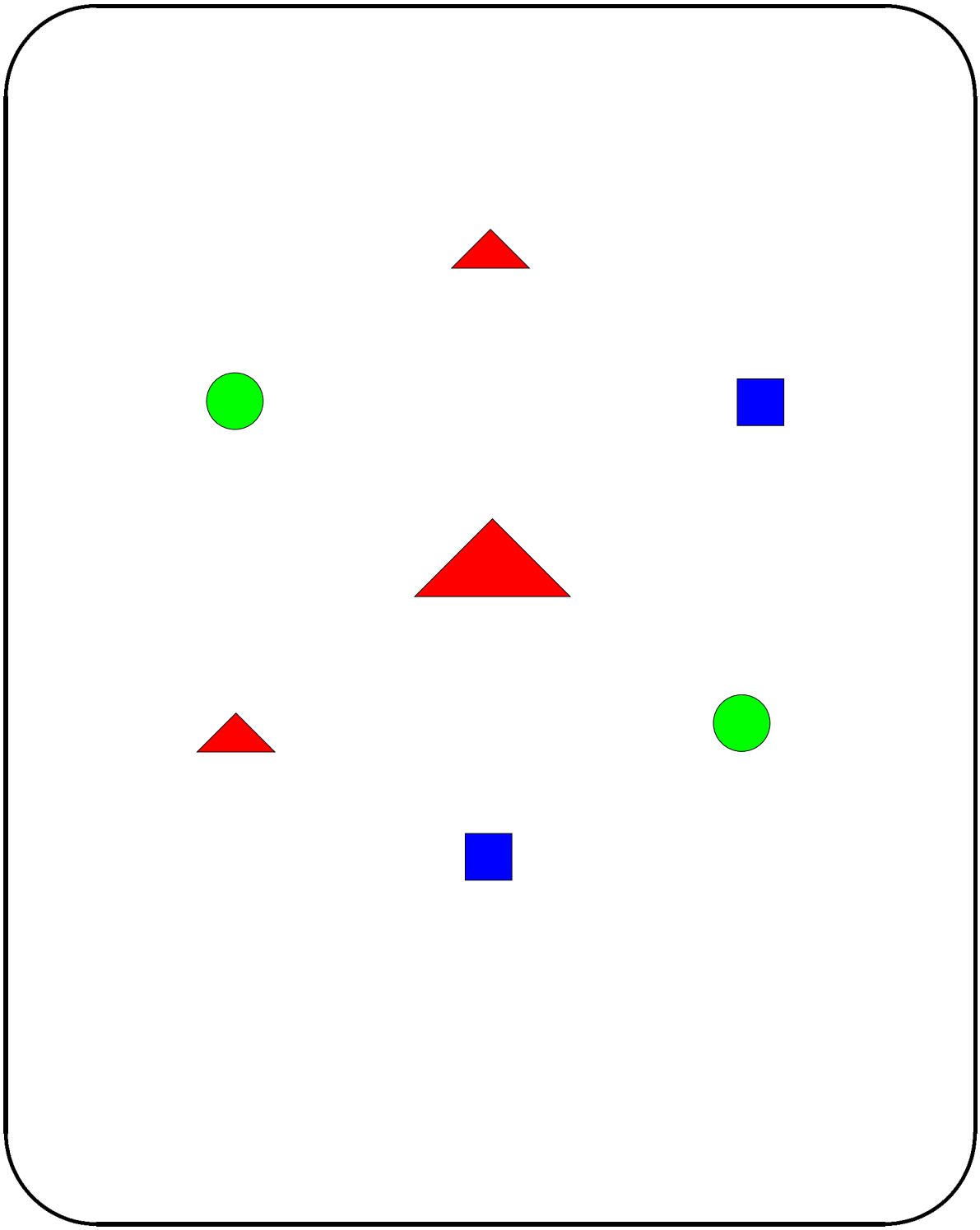


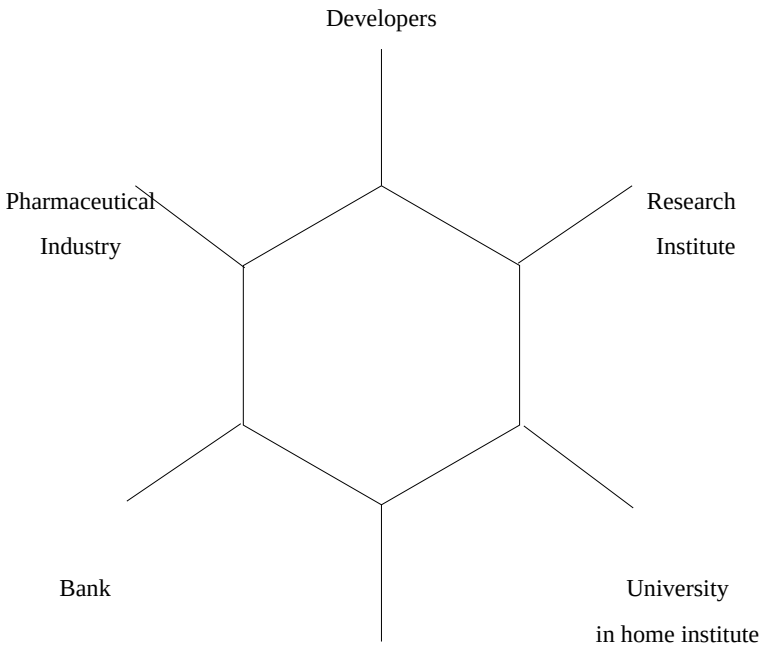

Student 1
Professor 1
Secretary
Server
Assistant Professor
Professor 2
Student 2

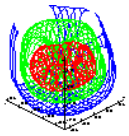

UNIX Machine
Personal Computer
Apple

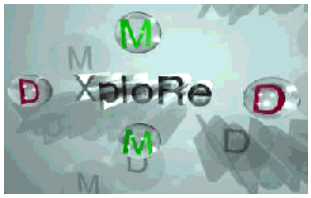



Student in Internet Cafe
Professor 1
in Vacations
Secretary
in home office
Assistant Professor
on research trip
Professor 2
in home institute
Child at Home (Berlin)




Developers
Pharmaceutical
Industry
Research
Institute
Bank
University
in home institute
Businessman
in Internet Cafe





combining statistical Methods and Data
the ideal platform for applied data analysis

Overview	Tutorials	Latest News
Download	Java interface	CGI interface
Help System	User Manual	FAQ

Search for keywords (help on searching) Search

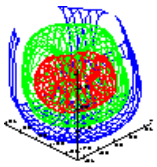
XploRe is a highly interactive computing environment. It combines the application of statistical procedures as well as the development of algorithms for statistical exploration of data.

The main features of XploRe are:

- fast programming language
- structured data types
- hierarchical help system
- graphical user interface
- network facilities
- autopilot helpsystem
- remote procedures calls
- dynamic link calls

XploRe is available for

- Sparc Solaris 2
- Linux 2 (Elf)
- HP-UX
- DEC Alpha
- SGI

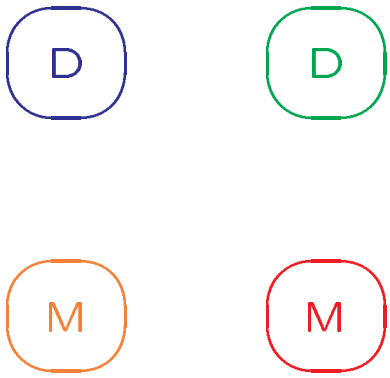


statistical Data

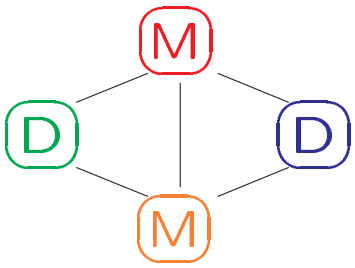
and

statistical Methodology

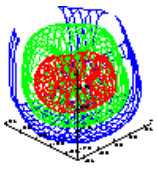
are on the net !



statistical practice before



statistical practice now



Scenario 1

Dataset of socio economic panel (SOEP) on migration East $\rightarrow$ West

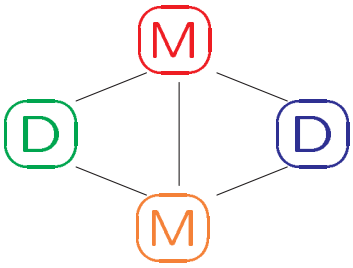
Dataset of actifs financiers (INSEE) on the savings behavior

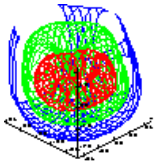
Statistical Method: probit analysis in a certain statistical system

$$P(Y = 1|x, t) =$$

Probit	Linear-additive
$G(x^T\beta_1 + t^T\beta_2)$	$G\{x^T\beta_1 + \sum_j g_j(t_j)\}$

Statistical Method: single index model or linear additive probit analysis





Scenario 2

α) We learn about a new statistical technology (e.g. linear-additive)

β) We program it in the ..... system

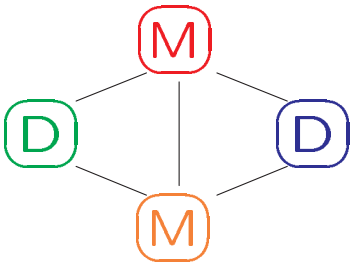
< time flies by >

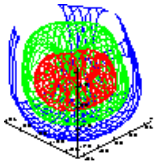
ε) Let me show this new method to my new employer!

< tempus fugit >

σ) It takes me so much time to understand what I have programmed...

ω) Why was there no networked help system?





Scenario 3

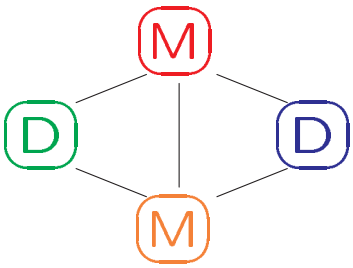
Marketing of faculty technology

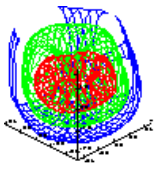
HHTest	Trellis' Display
ADEDis	Schafgans
HHMult	WIAS method

Let people in the world try your methods!

today:	Webpage static!
tomorrow:	NetStat Technology is useable
future:	MD*Tech

Method and Data Technology Center





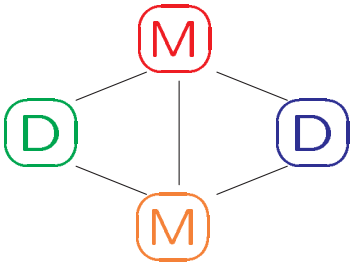
Scenario 4

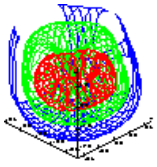
Decentralized teachware

- Course outline
- Examples
- Exercises

http://wotan.wiwi.hu-berlin.de

- Statistics
- Courses (password protected)

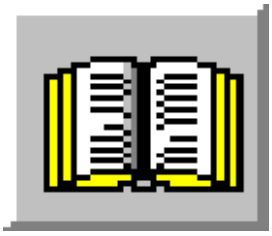




Scenario 5

Dynamic text books

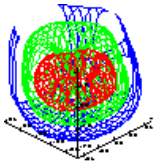
A text book is often seen as too dry information.



book

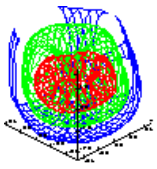
computer

tutorial



Hardware & Software bricks

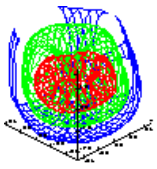
- Unix and/or Windows based systems
- Motif (GUI)
- HTML technology
- browser (Netscape, Explorer)
- background processing
- Java



The help system

Demands

- keyword oriented
- structured, but not too many layers
- no stiff hierarchy (e.g., tree)
- links to related information
- different instances (short, detailed, printed, etc.)
- accessible on the net
- provides executable examples
- includes user written procedures

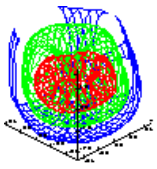


Implementation

The help system

HTML meets the requirements:

- text attributes (bold face, paragraphs, etc)
- links between keywords and documents
- icons or pictures can be included
- links can be considered as graph $\Rightarrow$ flexible hierarchy
- document can be distributed over the net

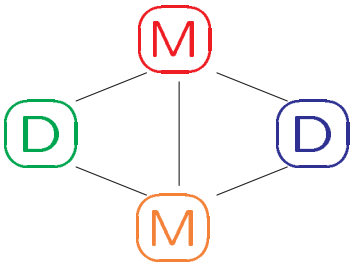


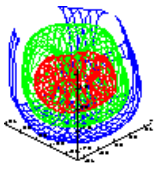
Implementation

Generate the help text not only in HTML:

short help	describes usage of a function or procedure
printed help	can not be substituted by online help

Translate the base help into different formats: HTML, short help (e.g., ASCII) and a word processor format (e.g., LaTeX)



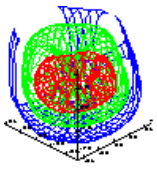


Network interface

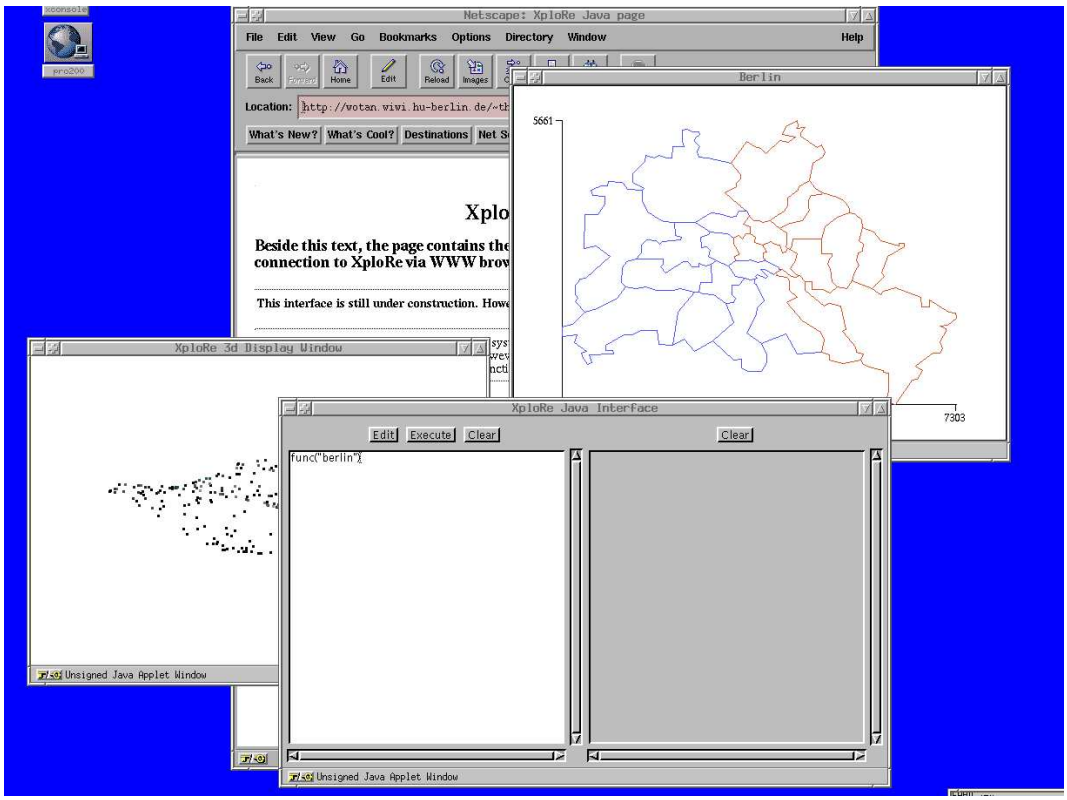
Java interface using a Java-capable browser

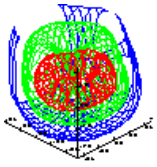
1. load the start page
2. Java applet starts and connects to the server
3. user dialog goes over the Java interface
4. displaying of text and graphical results
5. graphical results can be controlled locally

Advantages: No problems with the session management. Remote computing and local user interface.

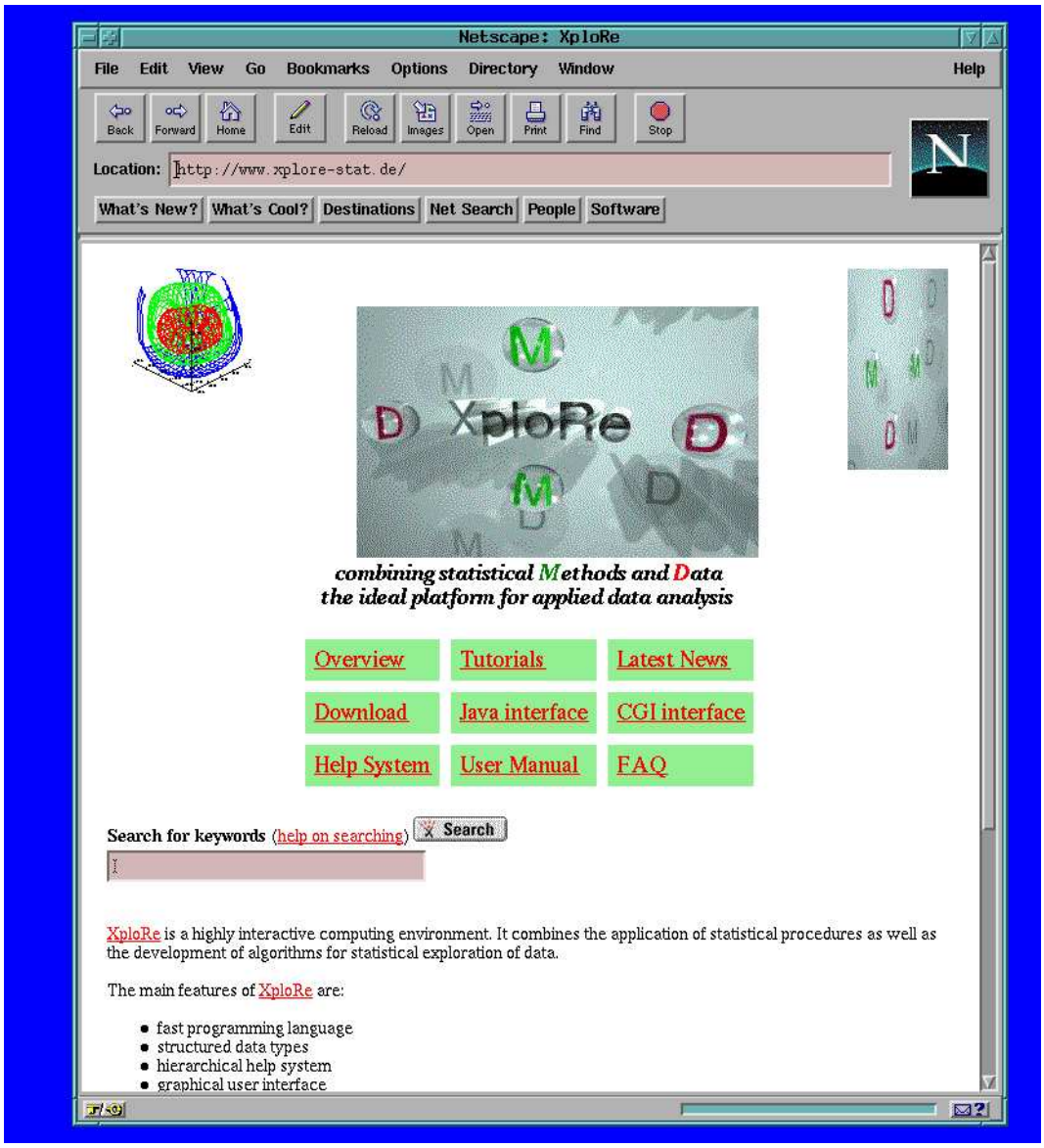


JAVA compatible browser

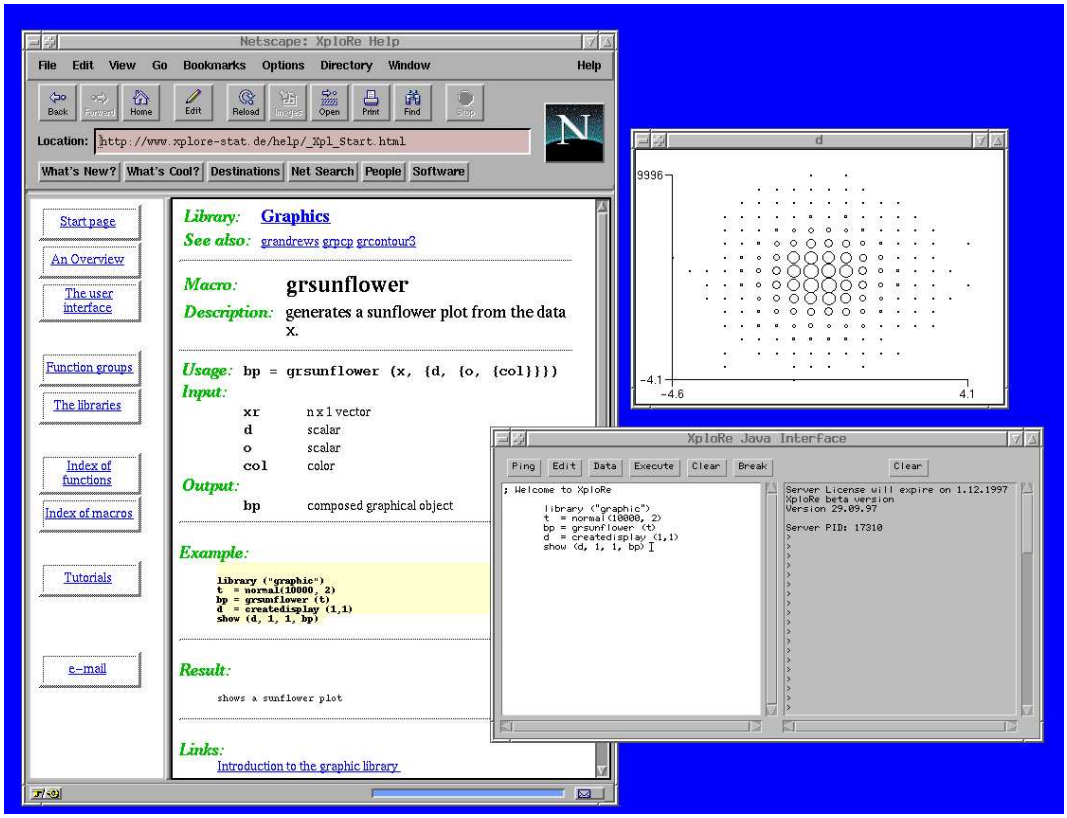


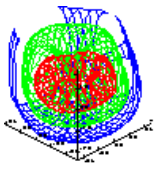


XploRe

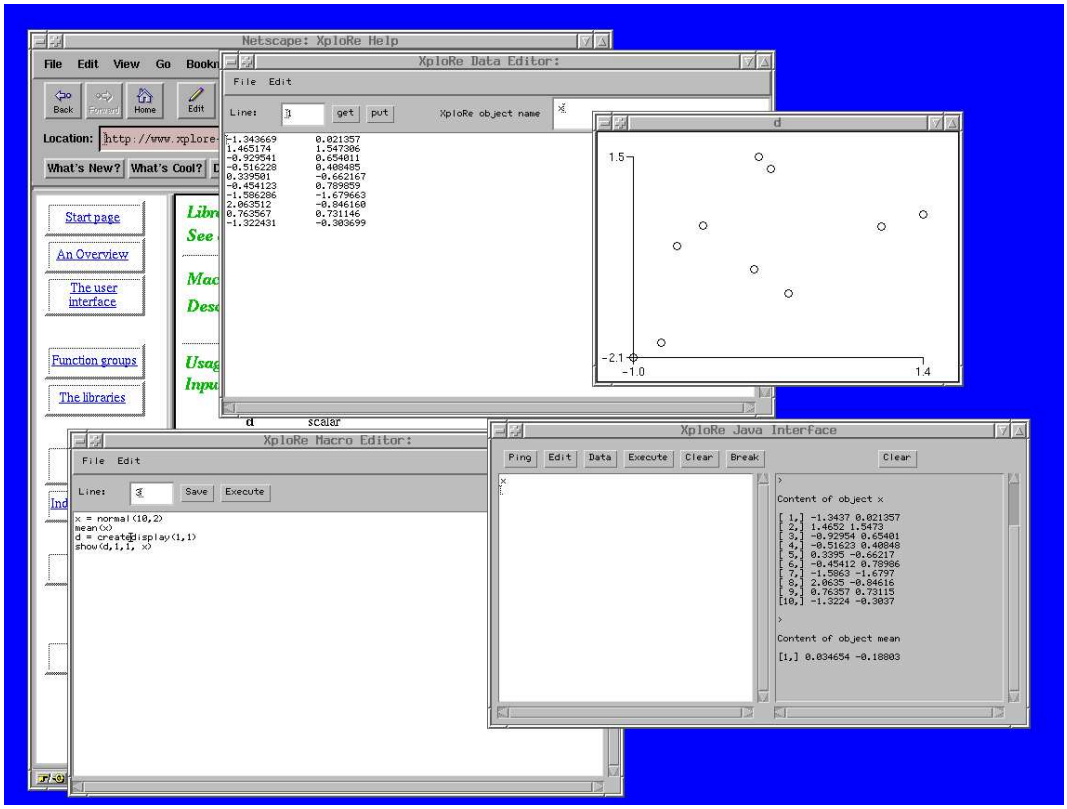

Netscape: XploRe
File Edit View Go Bookmarks Options Directory Window Help
Back Forward Home Edit Reload Images Open Print Find Stop
Location: http://www.xplore-stat.de/
What's New? What's Cool? Destinations Net Search People Software
XploRe
combining statistical Methods and Data
the ideal platform for applied data analysis
Overview
Tutorials
Latest News
Download
Java interface
CGI interface
Help System
User Manual
FAQ
Search for keywords (help on searching)
Search
XploRe is a highly interactive computing environment. It combines the application of statistical procedures as well as the development of algorithms for statistical exploration of data.
The main features of XploRe are:
fast programming language
structured data types
hierarchical help system
graphical user interface

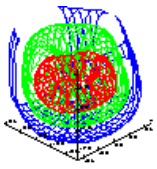
JAVA Interface 1



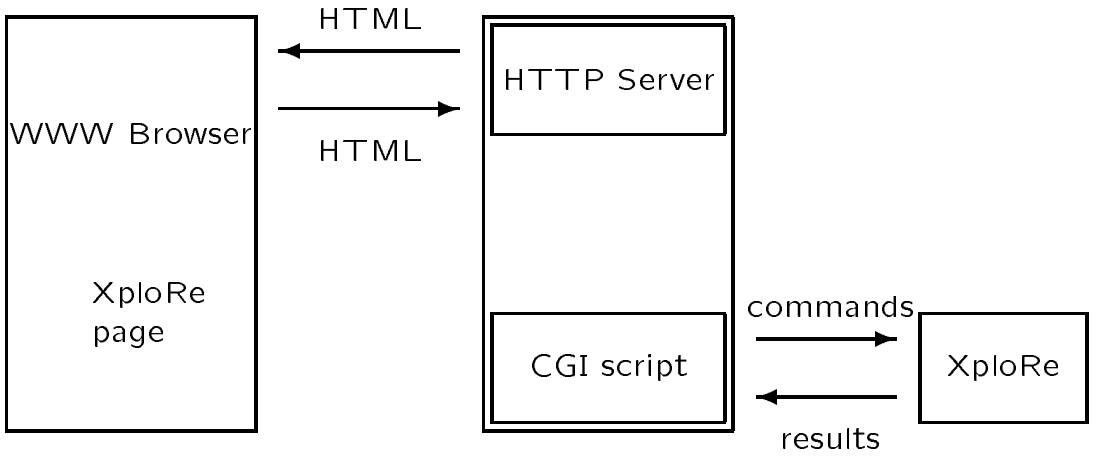


JAVA Interface 2

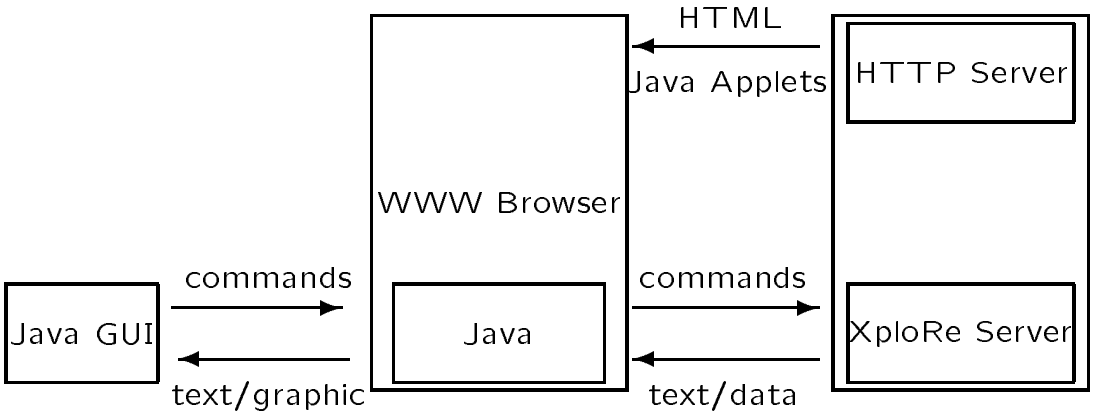


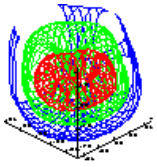


CGI-based interface



Java interface

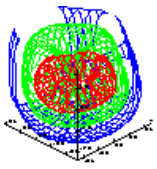




Problems of the CGI-XploRe interface:

- connection document based
- graphical representation
- no interaction
- no status handling

They could be solved (translation into a bitmap format, a more sophisticated server), but performance and handling would still be poor.

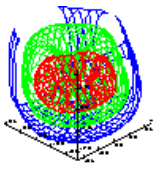


Implementation:

- HTTP-server starts non-graphic version of XploRe
- passes commands of the form
- results are wrapped into HTML
- XploRe computes the results and finishes
- send back to browser

Advantage:

- no restriction for client/server communication
- client can handle interaction locally
- all elements of the user interface can be implemented



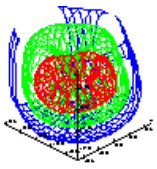
Problems by WWW access
(due to security mechanisms):

- XploRe- and HTTP-server on the same machine
- local file I/O not permitted

These problems do not exist for local installed XploRe-Java classes !

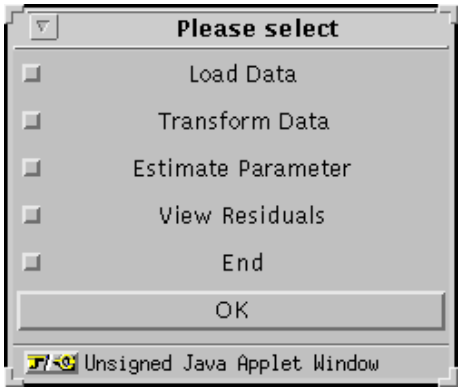
Extensions

- Improve server installation
- implementation of all user interface elements
- local file I/O



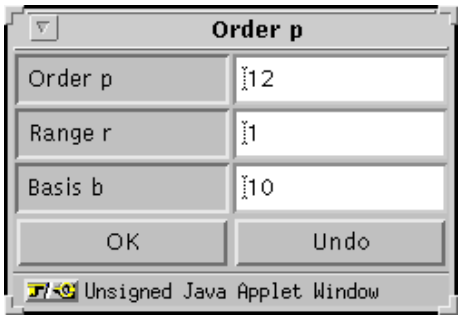
Graphical user interface:

selection box

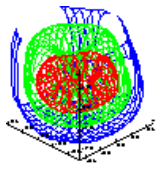


serves as interactive flow control

input box



serves as interactive parameter input



Application of Java interface

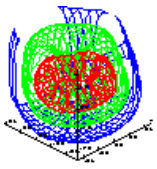
Data:

quarterly seasonal adjusted for West Germany, logarithm of first difference

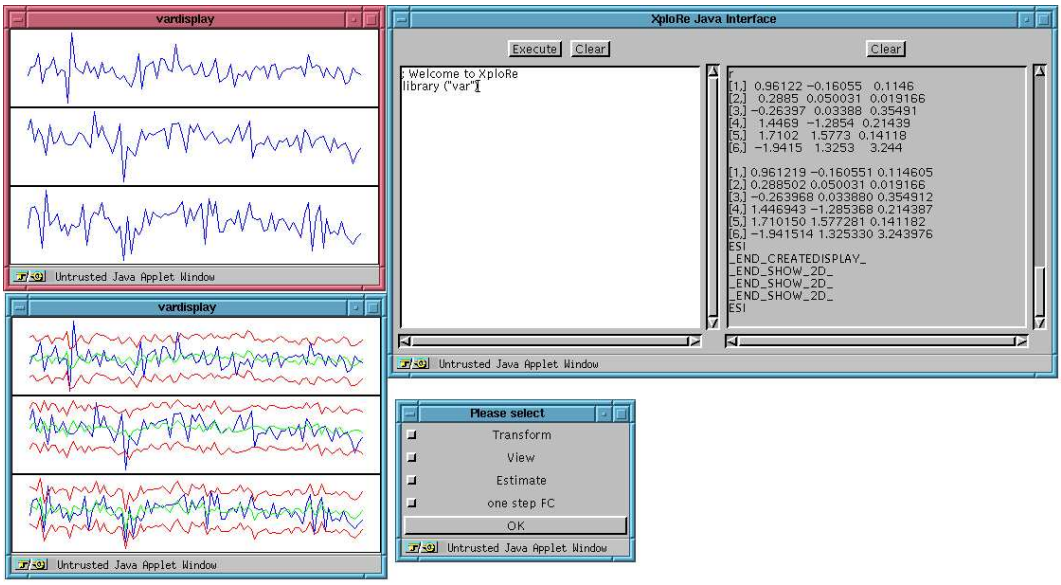
1. fixed investment,
2. disposal income
3. consumption expenditures

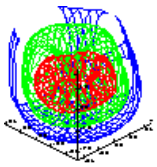
VAR:

$$y_t = \nu_t + \sum_{i=1}^{p} A_i y_{t-p} + u_t$$



Estimation and forecast of a vector autoregressive VAR(1) model:





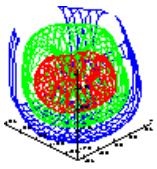
Hard- and Software

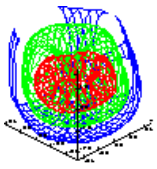
Three instances of XploRe

- text based (graphics only with PostScript documents), batch version
- GUI based (X11/Motif and later Windows NT/95)
- network based either on plain HTML or Java (WWW browser)

Implementation using lex, yacc, C, C++, Java, X11, Motif, Windows NT/95

Operating systems: Linux, SPARC, Windows, HP/UX, IBM AIX, SGI.



What to do to become a MD*Tech Center for 1 month?

1. http://www.xplore-stat.de
2. download Client/Server
3. download the JRE from http://www.javasoft.com
4. convert Your Tech into HTML
5. market Your Tech