GeoJabber

Finding Significant Analytic Events in Collaborative Visual Analysis Sessions

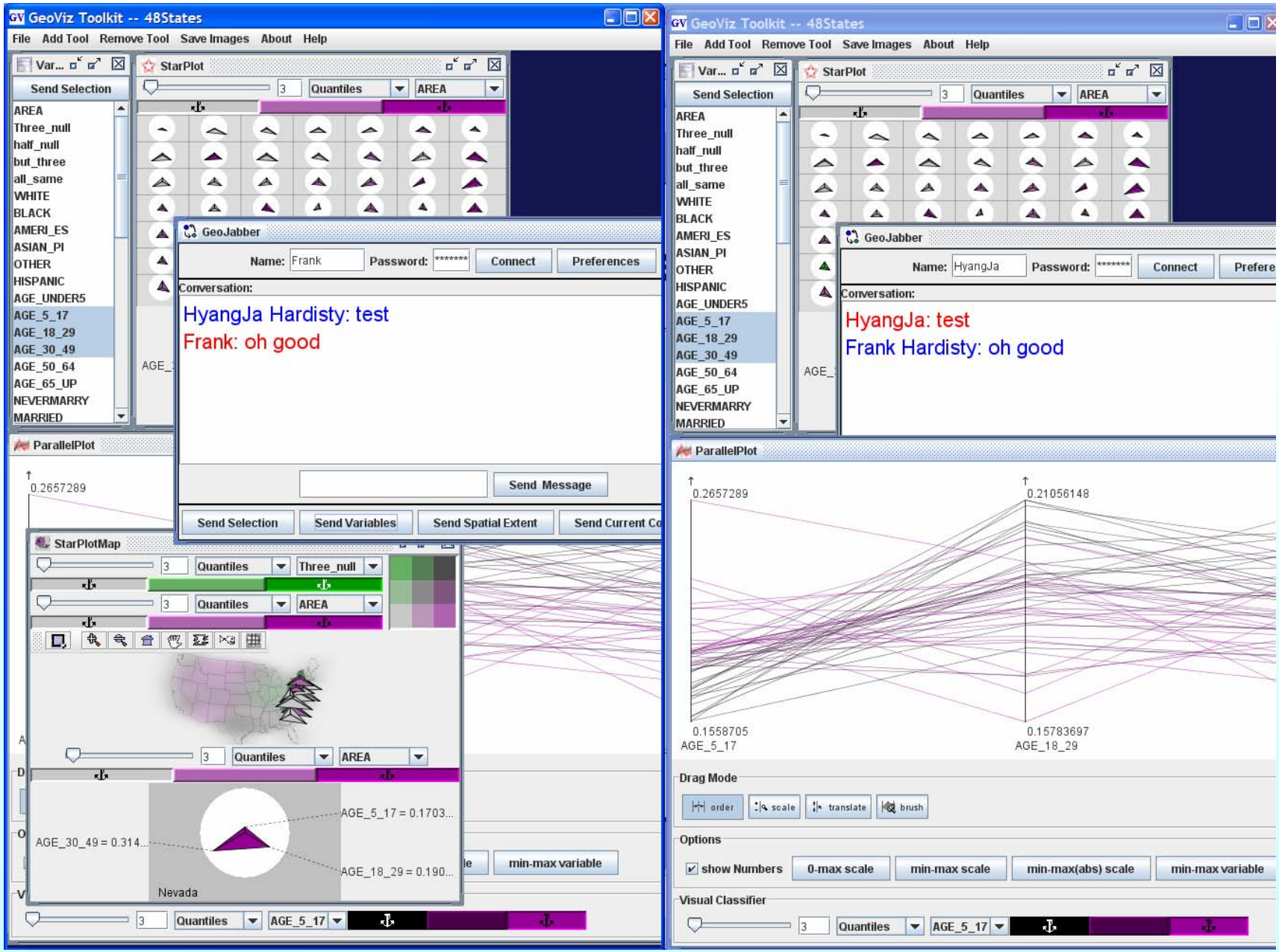
Frank Hardisty
GeoVISTA Center
Dutton e-Education Institute

Outline

- Demo
- Background
- Research Objectives
- Methodology
- Conclusions – Future Work

Demo

Demo – Dessert Before Dinner!

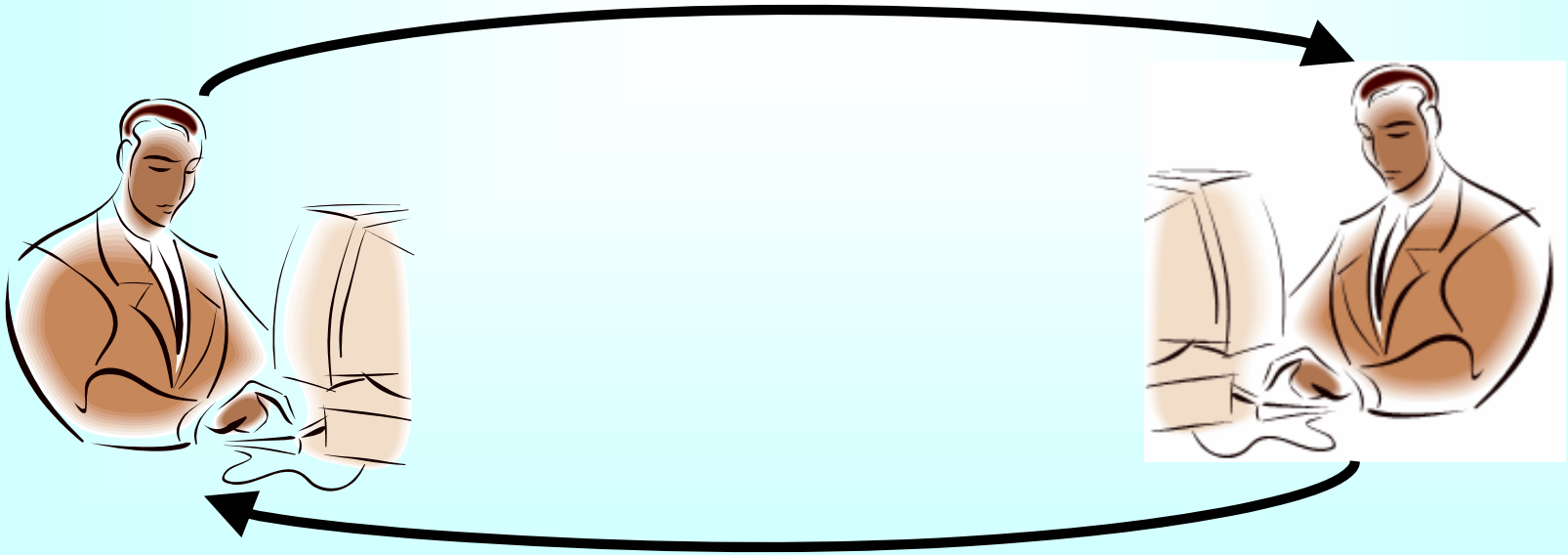


Background

Background

- **Scientific research using geospatial data is usually carried out by teams.**
- **Software has often been a limiting factor in collaboration.**
- **We need to extend geovisual analysis software to enable geographic collaboration over the network**

- **Type of collaboration being investigated – Different place, same time collaboration – For example, people coordinating response to a flood or other emergency**



Research Objectives

Research Objectives

- **Communication**
- **Typology**
- **Persistence**
- **User Analysis**
- **Technology Choices**

Communication

- **Provide a communication channel inside geospatial visual analytics software, supporting synchronous contact, including message sharing and knowledge sharing.**

Typology

- **Provide a typology of what should be shared between collaborators.**

Persistence

- **Provide a means of making analysis sessions and analysis artifacts persistent and accessible.**

User Analysis

- **Explore how user-entered text (chat) can serve as markers to find significant points in an analysis session.**

Technology Choices

- **Understand how technology choices affect human-human collaboration when using geospatially enabled software.**

Methodology

Methodology

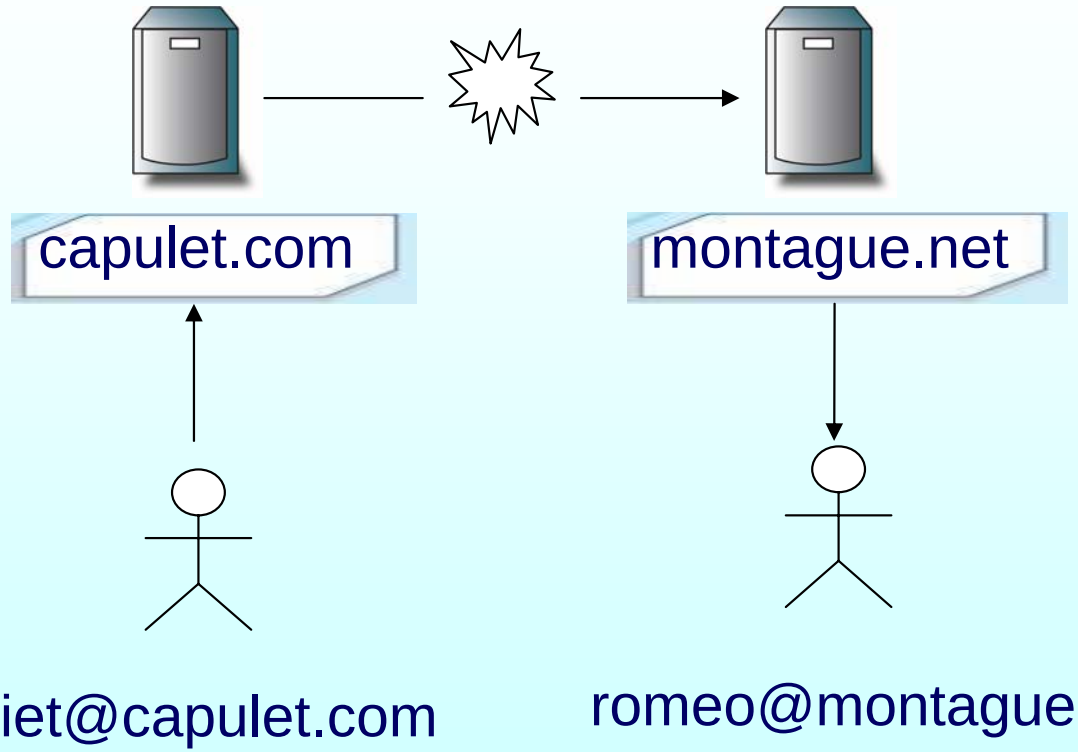
- Communication
- Typology
- Persistence
- User Analysis
- Technology Choices

Communication – Jabber

- Jabber is an open Instant Messaging (IM) protocol, like closed protocols defined for AOL's AIM, ICQ, Yahoo and Microsoft's MSN Messenger. Google Chat uses Jabber.
- There are a number of free and open source client and server Jabber programs available.
- Jabber messages are based on XML, and so can be used to pass Java objects or geographic coordinates

Communication – Jabber

- **Sample XML Jabber message from Juliet to Romeo**



Communication – Jabber

- Sample XML Jabber message from Juliet to Romeo

```
<message from='juliet@capulet.com'
         to='romeo@montague.net'
         id='message22'>
       <body>
         Wherefore art thou, Romeo?
       </body>
</message>
```

Communication – GeoJabber

Sample spatial extent message in GeoJabber

```
<message id="V2HVe-5" to="hyangja">
      <body>extension</body>
      <SpatialExtent xmlns="geoviz">
            <width>902.049560546875</width>
            <height>558.5320739746094</height>
            <y>-289.28460693359375</y>
            <x>-316.7908935546875</x>
      </SpatialExtent>
</message>
```

Communication – Marshal Software Objects to XML

- **Marshaling java objects converts them to XML. Marshaling is roughly equivalent to serialization.**
- **XML and Java are a good (not great) fit.**
- **This strategy leverages component-oriented and event-driven software development as seen in GeoVISTA Studio and the GeoViz Toolkit**

Typology

- **Partially Implemented**
- **What visual analytic states should we pass over the wire as XML?**
  - **Data**
  - **Display**
  - **Category**

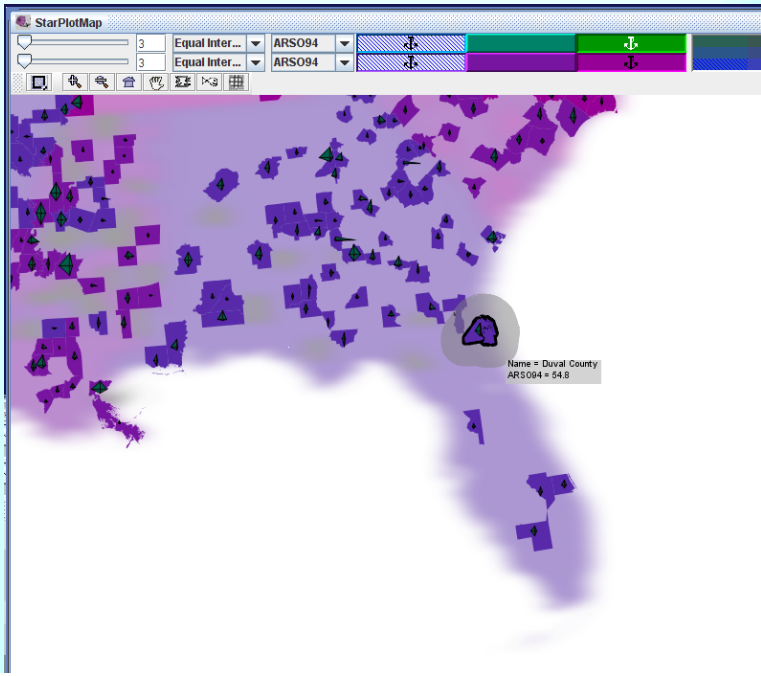
Typology - Data

- **The units of analysis**
  - **Spatial data**
  - **Temporal data**
  - **Numerical data**
  - **Textual Data**
  - **Derived data (via calculation)**

SUM_MAORI	SUM_PACIFI
450	44
72	16
237	31
432	138
54	23
404	100
13	0
102	21
286	84
712	98
300	81
146	38
29	6
78	18
67	21
59	14
18	0
299	27
169	14
496	84

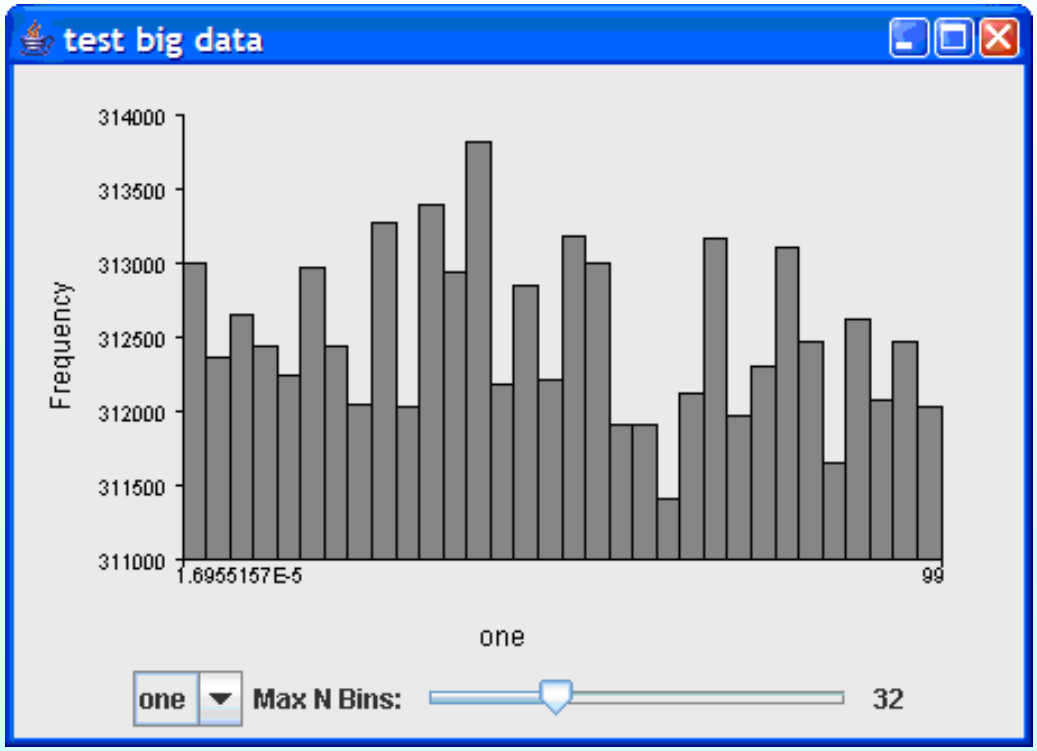
Typology - Display

- **How data is represented visually**
  - **Static properties**
    - **Color, location, size,**
    - **Typologies from MacEachren, Wilkenson, Bertin, others**
  - **Dynamic properties**
  - **Other sensory properties**



Typology - Category

- **Defining sets on data**
  - **Classifying**
    - **Univariate**
      - **Quantile**
      - **Equal Interval**
    - **Multivariate**
      - **K-Means**
      - **SOM**
  - **Extent**
    - **Attribute**
    - **Spatial**
    - **Temporal**



Persistence

- **Partially Implemented**
- **We can persist the program state, or artifacts of the program state, such as visual representations, and textual annotations**
- **Since the program state is preserved as XML, we can think of many way it can be stored, used, and accessed**
  - **On local disk**
  - **On a remote site**
  - **In a database**

User Analysis

- **Implementation begun, but not functional yet.**
- **HCI (Human-Computer Interface) researchers have focused on examining patterns in user behavior.**
- **One shortcoming of this work is determining where a significant event in the analysis took place (what does insight look like?)**
- **GeoJabber has the potential to help identify significant events in analytic sessions.**

Technology Choices

- **For Jabber -> OpenFire and Smack**

- **For marshaling -> XStream**

The alternatives…

Technology Choices

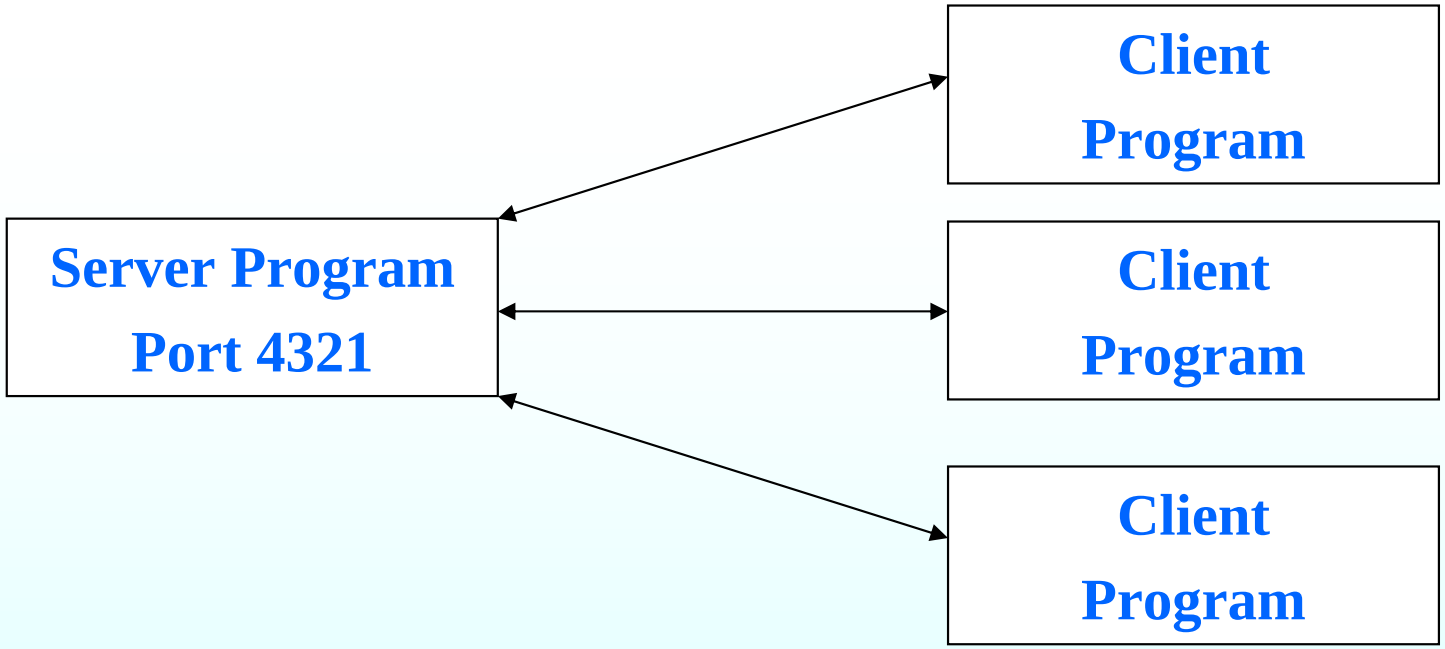
- Alternative Implementation Options for Collaboration
  - Sockets
  - JXTA
  - RMI
  - Jabber
- Alternative Options for Marshaling

Alternative – Sockets

- **Sockets let you send raw streams of bytes back and forth between two computers, giving you low-level access to the TCP/IP or UDP protocols.**

- **Java Sockets are a mechanism for communication built into most versions of Java.**

Alternative – Sockets



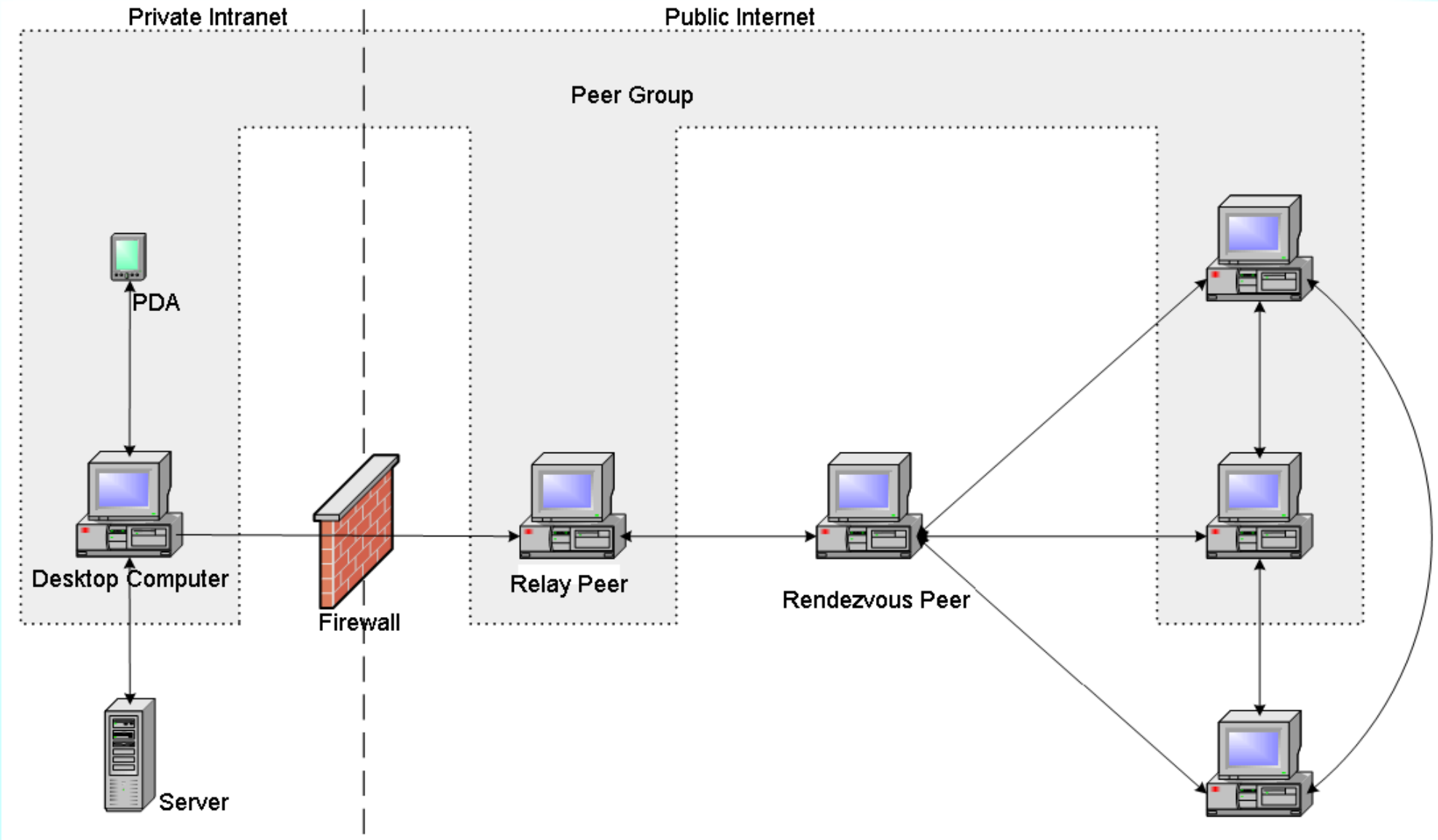
- **Disadvantage – sockets are too low level, and therefore do not support the concept of objects.**

Alternative – JXTA

- **A set of open peer-to-peer protocols.**
- **Allow any connected device on the network to communicate and collaborate in a P2P manner.**
- **Create a virtual network where any peer can interact with other peers and resources directly.**

Alternative – JXTA

- **JXTA Network Overview**



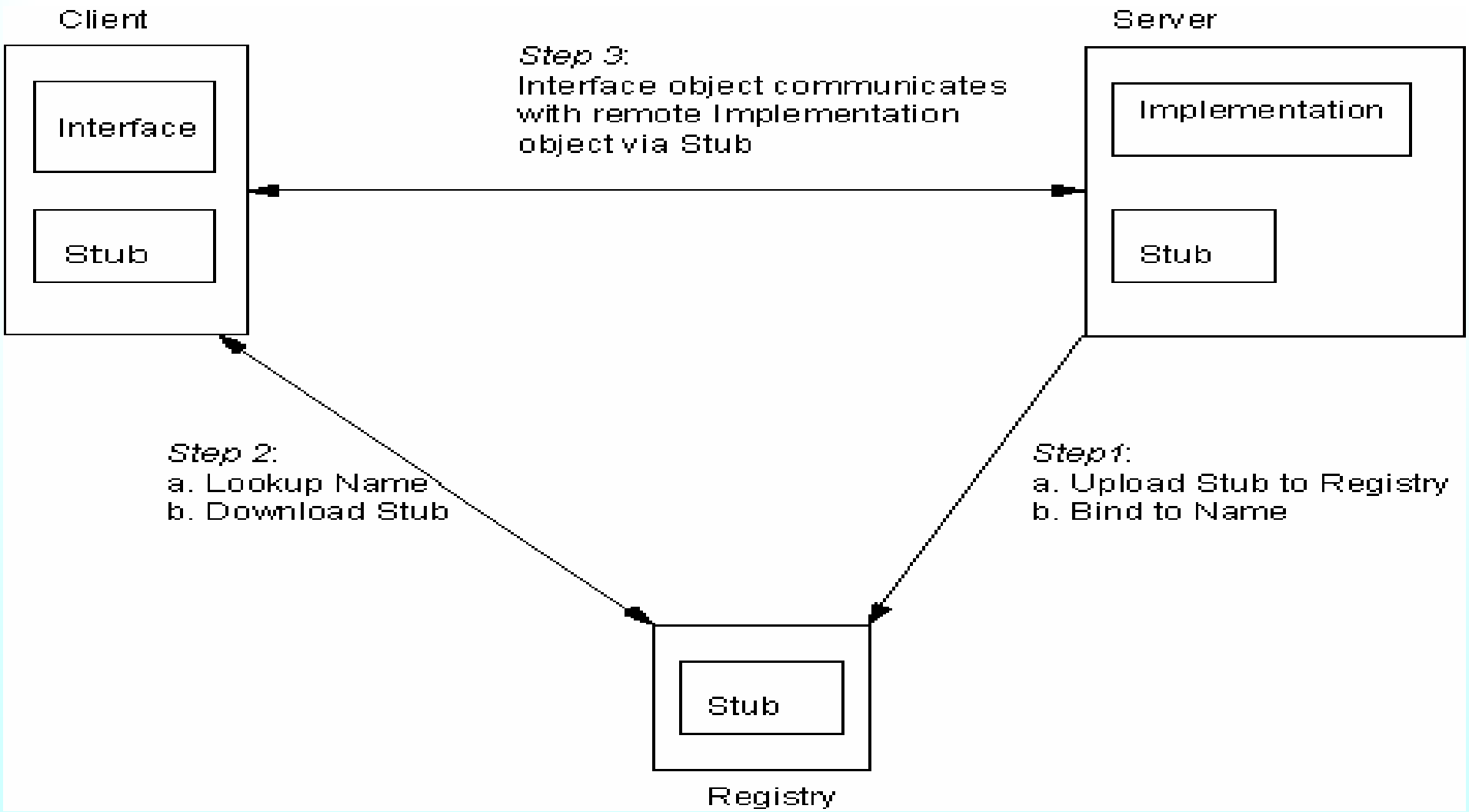
Alternative – JXTA

- **Disadvantage – complicated and unfamiliar. Did not work out of the box (at least for me). Because JXTA does not use TCP/IP or UDP, programmers need to learn an entirely new set of networking protocols to debug JXTA networking problems.**

Alternative – RMI

- **Remote Method Invocation (RMI)**
- **An object running in one Java Virtual Machine (VM) invokes methods on an object running in another Java VM.**
- **Remote communication between programs written in the Java programming language.**

Alternative – RMI



Alternative – RMI

- **We selected this solution for our first implementation for ease of use.**
- **RMI is part of the standard Java distribution**

Alternative – RMI

- **Disadvantages of RMI**
  - **Fragile against change. Versions must be kept tightly in synch. If one client updates and another does not, the application will fail.**
  - **For peer-to-peer applications, each node must be a client and a server, increasing developmental effort.**

Alternative Jabber Implementations

- **Tigase – Java. Easier to extend, but harder to get started.**
- **eJabberd – Written in Erlang. High concurrency.**
- **Many more**
- **Cisco bought Jabber, inc. for no reason I can tell.**

Alternatives to XStream

- JAXB
- XMLBeans
- Many others….

Conclusions – Future Work

Conclusions – Future Work

- **Conclusions**
- **Future research**
- **Acknowledgements**

Conclusions

- **GeoJabber provides a working method for allowing synchronous collaboration in Visual Analytics.**
- **GeoJabber is extensible, both in what types of analytic operations are supported, and what is done with the analysis artifacts.**
- **Much work to be done….**

Future Research

- Support more types in my typology.
- Integrate with annotation facilities.
- Provide search over previous sessions.
- Publish to G-EX portal
- Publish to Penn State Annotations DB
- In the misty future, provide interoperation with ArcGIS, Open Layers, etc…

Acknowledgements

The GeoVISTA Center

NEVAC (North-East Visualization & Analytics Center)

However, any opinions, findings, and conclusions or recommendations in this document are those of the authors and do not necessarily reflect views of the U.S. Department of Homeland Security.

Workshop Organizers