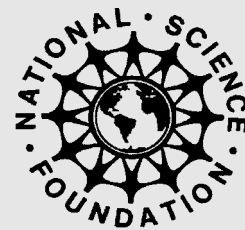
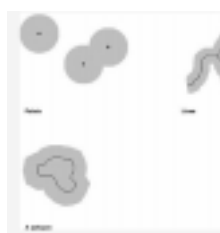
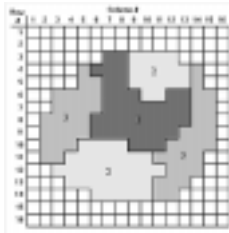
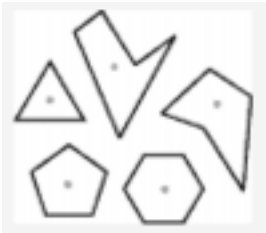


**Citizen Science  
Toolkit Conference**

June 20 - 23, 2007

projects and problems:  
location, scale and precision in  
citizen science

Dave Lemberg  
Director, Michigan Heritage Water Trails  
Associate Professor, Department of Geography  
Western Michigan University



CORNELL LAB of  
ORNITHOLOGY

CORNELL LAB OF ORNITHOLOGY

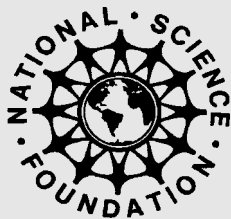
607.254.BIRD telephone  
[www.birds.cornell.edu](http://www.birds.cornell.edu)

159 Sapsucker Woods Road  
Ithaca, New York 14850

This presentation took place at the Citizen Science Toolkit Conference at the Cornell Lab of Ornithology in Ithaca, New York on June 20-23, 2007.

Note that this document did not originate as a formal paper. Rather, it combines an oral presentation with accompanying PowerPoint slides and reflects the more informal, idiosyncratic nature of a delivery prepared specifically for this live event.

Documentation of the conference is meant to serve as a resource for those who attended and for others in the field. It does not necessarily reflect the views of the Cornell Lab of Ornithology or individual symposium participants.



This documentation is supported by the **National Science Foundation** under Grant ESI-0610363.

Any opinions, findings, and conclusions or recommendations expressed in this documentation are those of the authors and do not necessarily reflect the views of the National Science Foundation.

---

The following is one of three focus point presentations delivered as part of the session titled "Technology and Cyberinfrastructure" on day two of the Citizen Science Toolkit Conference

---

For complete documentation of conference proceedings and to learn more about citizen science and the Citizen Science Toolkit, or to join the ongoing citizen science community, go to:

**<http://www.citizenscience.org>**

# Projects and Problems: Location, Scale, and Precision in Citizen Science

## Introduction

This is going to be a big change of pace as well as a change of title. While the original title was “Mashups for Michigan Heritage Water Trails,” I really don’t have a lot to say about mashups and heritage water trails, though I do want you to know what a heritage water trail is. Really what I am going to be talking about is geography; about projects and problems of location, scale, and precision in citizen science.

Dave Lemberg,  
Director,  
Michigan Heritage Water Trails;  
Associate Professor,  
Department of Geography,  
Western Michigan University



I don't do a lot of  
citizen science,  
yet,  
but I'd like to.

I don't do a whole lot of citizen science, yet, but I'd like to. Janis came to Western Michigan and talked a lot about the Bluebird project, but also talked about citizen science. I thought, wow, this is a really cool thing for what I'm doing. What I do is work with communities and citizens on

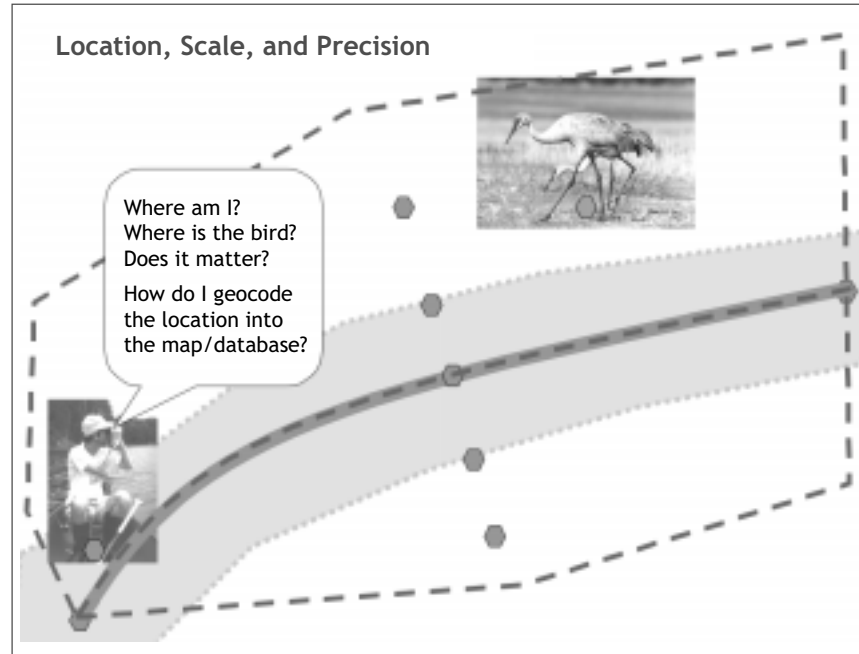
planning problems in various ways, economic development mostly. I applaud Sarah Kirn for what she is doing at the Gulf of Maine Institute in working with fishermen to connect research with economic development and jobs. That's what I do, dealing with things like ecotourism, looking at various planning problems, and I'll talk a little bit about those. But what we are really going to be focusing on is location.

I am the son of a real estate broker, and what did I hear from day one? Location, location, location. Location is everything, and location is really an interesting problem in terms of how we encode it. I'm floating down my river and I see a Sandhill Crane (fortunately we're seeing a lot of Sandhill Cranes, they're all over the place in southern Michigan now).



[www.wmich.edu/glcms/watertrails](http://www.wmich.edu/glcms/watertrails)

Where do I place that crane? Do I place it where I am, where the crane is, or someplace on the river, on either side of the river, in a point in polygon? That is more or less what we will be focusing on here.



## Projects, Problems, Issues

### The Michigan Heritage Water Trails

Now I'll tell you a little about some projects. I'm Director of the Michigan Heritage Water Trails program. Let me give you a little

history. Michigan was one of the last states to get a water trails program. Water trails are trails on rivers and lakes that you paddle by canoe or kayak, mainly kayak nowadays. We couldn't get a bill through the legislature getting a state water trails program because the state is broke (because of Michigan and the car industry) and they

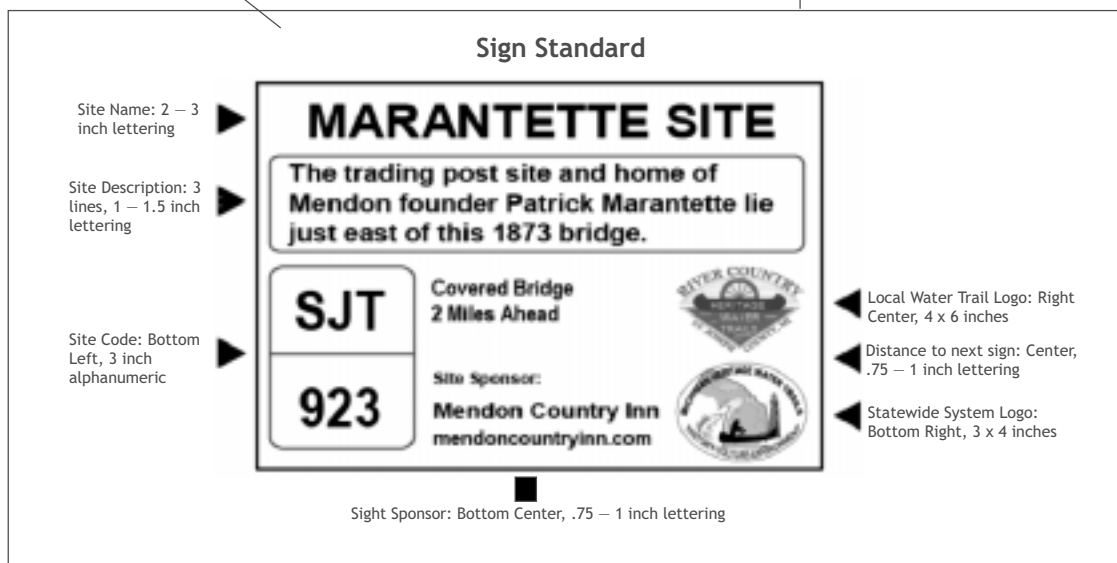


didn't want to fund it under a government recreation program.

So we repackaged it. The Michigan Heritage Water Trails program is an economic development and education program. The basic idea is that we map the river and we do research on history, culture, and the environment, and then we interpret it. As you paddle, you see the signs, you look at your interpretive guide, and you learn as you go down the river.

One of our problems is that we'd like to get people back on these trails after doing them once, and we want people to go to these villages and spend money. So how do we get them back? Well, maybe we can get them back if they start monitoring the river. If they do it on a regular basis, they're going to eat at the restaurants along the way regularly as well.

This is our basic sign format: the name of the place, a little bit of interpretation to tease you, logos, and most importantly, down at the bottom, is the sponsorship.



The entire thing is funded by sponsorships. You get somebody to pay \$500 for the sign sponsorship and the sign only costs about \$35. The rest of it goes to print the interpretive guide, which we then sell, which then gives you money to implement the trail and go on to the next.

What we are doing with mashups is that if you want to know where to go paddling, you can punch in your address and say how far you want to paddle and how far you want to drive, and the program will tell you what trails are available and what regions of the river are suitable for what you want to do, and we're filtering that system through a data mix. That product was then transferred to a project looking at street trees.



what polygon you have to be in. That is followed by a database query to a MySQL database, which is based on underlying grid cell formats. We also are doing online questionnaires with e-mail returns. Hopefully, we'll be switching into database for that.

### Spatial Data

Now let's switch over to geography. The spatial data that geographers deal with is based on space or location, theme or attributes, and time. That makes things wonderful for geographers because they deal with spatial data—space, theme, and time—so we can do anything we like. There is not much that is not spatial.

There are various ways of looking at space: address, coordinates, and looking at linear objects or polygon objects; thematic data—your nominal, ordinal, interval and ratio data, which then can be analyzed in different ways; and discrete and continuous time (or space for that matter).

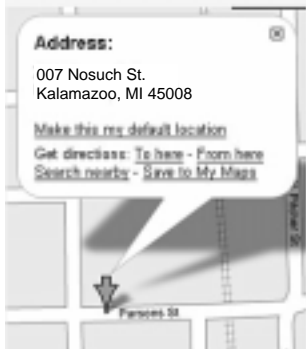
### Location Issues: Where is the Observation?

So, some issues: Where are you and how do you find it? Your street address? Latitude-longitude or some other coordinate system? Pointing on a map to where you think it is, more or less? Or using GPS, which may be accurate or not accurate depending on your equipment, the time of day, where the satellites are?


### Spatial Data

- Space (Location)
  - Address
  - Coordinates
  - Linear
  - Polygon
- Theme (Attributes)
  - Nominal
  - Ordinal
  - Interval
  - Ratio
- Time
  - Discrete
  - Continuous


### Point Data




Street Address



Latitude-Longitude



Map Location



GPS

You've got line data. Dealing with rivers, this system works fairly well. You know where you are between two points—the river reach, or the street segment. Sometimes you do milepost offsets, or centerline or left side/right side. I've done a lot of work on this once upon a time, doing a database standard for CalTrans, where every

regional group had a different database and interoperability was absolutely impossible.

### Line Data



Street Segment



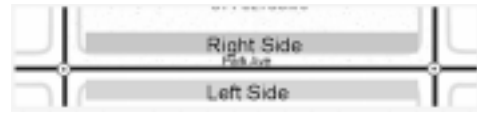
Milepost - Offset



River Reach



Centerline



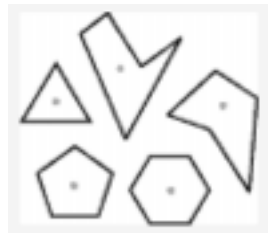
Right Side / Left Side

With polygon data you can look at bounded spaces; you can locate a point in the spatial center, a centroid; you can use grid cells; you can use buffers around points, around lines, around polygons or other minutia of where you are.

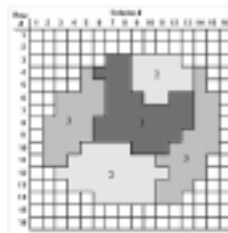
### Polygon Data



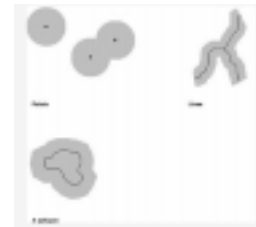
Bounded Spaces (polygons)



Centroids



Raster (grid cells)



Buffers (point, line, polygon)

### Scale Issues

One of the issues of scale is in terms of your points: Are you looking at the spot where you are? Are you looking at some distance around where you can see a landscape scale? Are you looking at a region near the spot?

And it is the same with the line and polygon. Are you looking on the line? Are you looking at the corridor? Are you looking at line of sight from the line on either side, looking on each side from the river into the trees? Are you looking at some region around the line? Within polygons there are your grid cells; a parcel like your backyard; a



bounded area such as a county; or some unbounded area, a region, something that's somewhat fuzzy.

What is your neighborhood? Ask two different people what their neighborhood is and they'll give you two different answers. How do we determine how many birds there are in a neighborhood unless everybody's definition is the same? Think about it.

**What is the Scale of Observation?**

<ul style="list-style-type: none"> <li>• Point                             <ul style="list-style-type: none"> <li>- Spot location</li> <li>- Buffer (distance) around the spot</li> <li>- Line of sight from the spot (landscape scale)</li> <li>- Region near spot</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Line                             <ul style="list-style-type: none"> <li>- On the line</li> <li>- Corridor</li> <li>- Line of sight from line</li> <li>- Regional near the line</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Polygon                             <ul style="list-style-type: none"> <li>- Grid cell</li> <li>- Parcel (yard)</li> <li>- Bounded area</li> <li>- Unbounded area (region)</li> </ul> </li> </ul>
--	--	--

### Precision Issues

The big issue is, how precise do we need to be? We don't need to be right on the point for most data. Counting our birds, we need to know more or less where they are. For some analysis it may be important, but for other analysis, fuzzy is okay. So are we looking for accuracy or precision? How much effort do we want to put into precision if accurate is good enough? It's a "horseshoes or hand grenades" kind of thing: It's close enough if it works.

What are we observing and at what scale? Is it significant? Are there related data sets that we can compare with our data, and do we need to have interlocking scales? How much does location really matter for what we're observing?

And how precisely do the citizens locate themselves? Do they know where they are? How do they record or select their locations? We need to know that in order to develop our user interface. And do we really want to be precise? Do we want to hide our sites?

### Issues for Discussion

Below are some issues that I've talked about for later discussion. Are we allowing for appropriate types and classes of location? Are our current location entry and coding technologies adequate for what we are trying to do? How can we error check the locations that are being submitted? Are we thinking about appropriate scales of analysis, and are we thinking about appropriate scales of precision in our location schemes? Are we thinking about appropriate metrics, classes, and precision of our temporal data? A lot of times we think about space, but not so much about time. How can the technology be used to promote accuracy in our attribute data?

I have included some other related issues that I just thought were kind of cool. How can we integrate recreation and tourism with citizen science? I think there's a real future for that. And how can we

**How Precise Do We Need to Be with Our Location?**

- Is our objective accuracy or precision?
- What are we observing?
  - At what scale are our observations significant?
  - Are there related data sets with which we can compare our data?
  - How much does location matter for what we are observing?
- Do we want to be precise?
  - Do we want to avoid other visits to the site?

integrate citizen science with physical fitness and anti-obesity efforts, which are really important things? Looking around this room, it's pretty clear that it works. You look at a general population sample and the average obesity level is going to be much higher than it is in this group, and I applaud you all for that.

#### Discussion Issues

- Are we allowing for appropriate types and classes of location?
- Are our current location entry and coding technologies adequate?
- How can we error check the locations that are being submitted?
- Are we thinking about appropriate scales of analysis?
- Are we thinking about appropriate levels of precision in our location schemes?
- Are we thinking about appropriate metrics, classes, and precision of temporal data?
- How can the technology be used to promote accuracy in our attribute data?
- How can we integrate recreation and tourism with citizen science?
- How can we integrate citizen science with physical fitness and anti-obesity efforts?