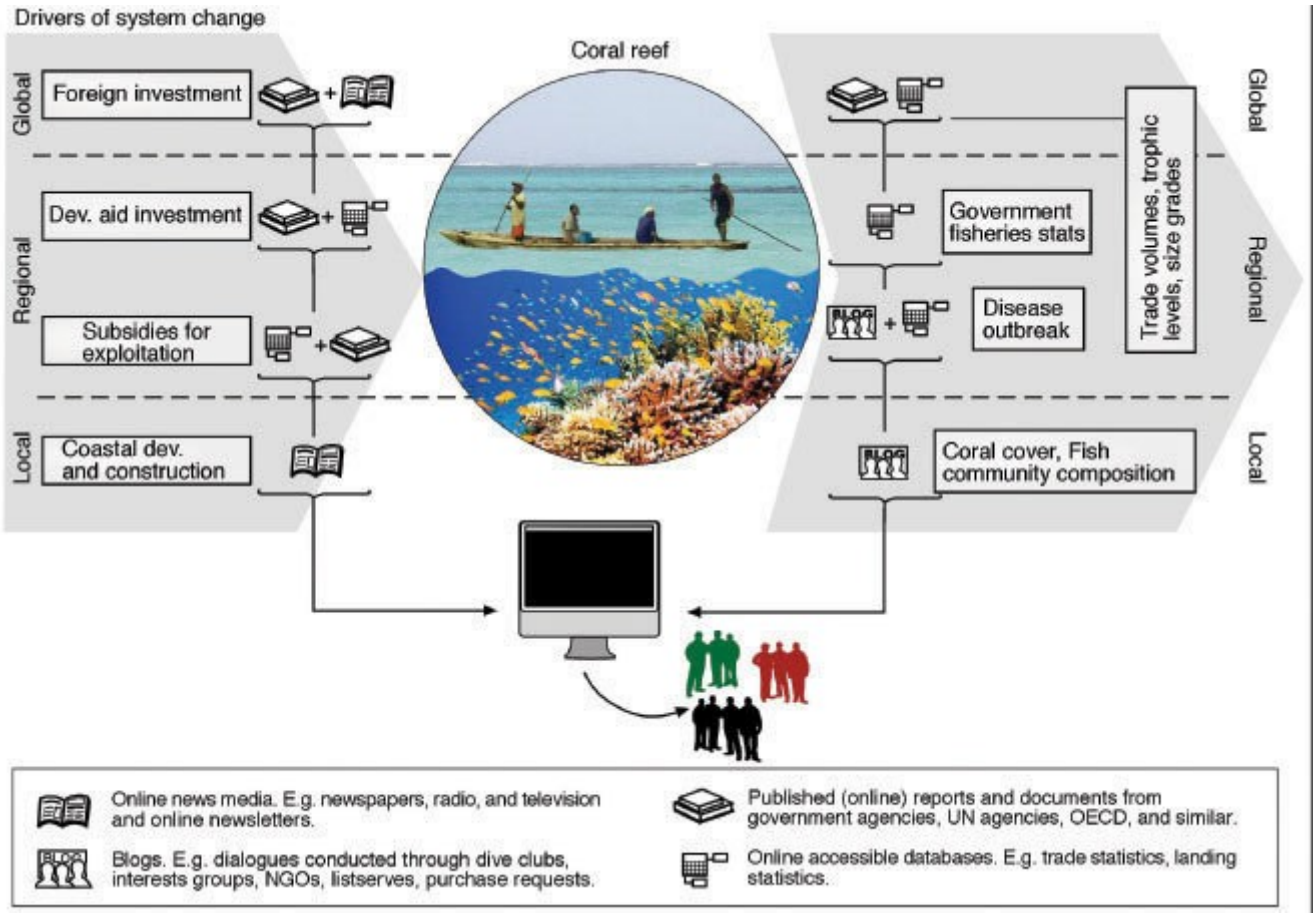
WIRED SCIENCE

Crawling the Web to Foretell Ecosystem Collapse

By Alexis Madrigal March 19, 2009 | 5:28:18 PMCategories: Environment



The Interwebs could become an early warning system for when the web of life is about to fray.

By trawling scientific list-serves, Chinese fish market websites, and local news sources, ecologists think they can use human beings as sensors by mining their communications.

"If we look at coral reefs, for example, the Internet may contain information that describes not only changes in the ecosystem, but also drivers of change, such as global seafood markets," said Tim Daw, an ecologist at the UK's University of East Anglia in a press release about his team's new paper in *Frontiers in Ecology and the Environment*.

The six billion people on Earth are changing the biosphere so quickly that traditional ecological methods can't keep up. Humans, though, are acute observers of their environments and bodies, so scientists are combing through the text and numbers on the Internet in hopes of extracting otherwise unavailable or expensive information. It's more crowd mining than crowd sourcing.

Much of the pioneering work in this type of Internet surveillance has come in the public health field, tracking disease. Google Flu Trends, which uses a cloud of keywords to determine how sick a population is, tracks epidemiological data from the Centers for Disease Control. Less serious projects — like this map of a United Kingdom snowstorm based on Tweets about snow — have also had some success tracking the real world.

These research efforts seem to indicate that people are good sensors, but pulling the information from what they post in human-readable formats and transforming it into quantitative models of the world is tough. The Global Public Health Intelligence Network has developed an epidemic warning system that pulls in data from news wires, web sites, and public health mailing lists. The GPHIN, which is probably the most advanced and uses highly variegated information, only picks up on about 40 percent of the 200 to 250 outbreaks that the World Health Organization investigates each year.

Nonetheless, Daw and and his co-authors from the Stockholm Univeristy Resilience Centre, say traditional ecological monitoring has its problems, too. Humans can make huge changes to ecosystems faster than the standard methods of data collection can keep up.

"The challenge is that existing monitoring systems are not at all in tune with the speed of social, economical and ecological changes," the researchers write on their blog.

By looking at human data, not just fisheries and ecological readings, they think they'll be able to detect ecosystem tipping points before they happen.

"Web crawlers can collect information on the drivers of ecosystem change, rather than the resultant ecological responses," they write. "For example, if rapidly emerging markets for high value species are known to be socio-economic drivers which lead to overexploitation and collapse of a fishery, web crawlers can be designed to collect information on rapid changes in prices, landings or investments."

But right now, their plans remain theoretical, and while scraping data seems easy enough, turning it into knowledge is another story. John Brownstein, a Harvard bioinformaticist and co-founder of HealthMap, which does for disease what Daw wants to do for ecology, said that applying the framework to ecology could work.

"There's no reason it can't be done," Brownstein said. "The only difference is that this is more difficult. The media and other sources are sensitive and fine-tuned to things like human disease. The threshold for the reporting of a mysterious disease is different from the threshold for an ecological phenomenon."

In other words, while reporters (or Tweeters) will include individual-level death data in human stories, massive die-offs or flora changes could very well go unnoticed and probably unquantified.

And even with disease data, there are serious signal-to-noise challenges. In a paper that Brownstein co-authored last week, he showed that monitoring search terms for disease indicators could have tipped officials off to a deadly outbreak of listeriosis in Canada. But spotting emergent diseases instead of ones that have already caused major damage is a more challenging proposition.

"It's so tough to figure out why people search for specific information," he said.

See Also:

- Scientists Mimic Beetle's Liquid Cannon
- Mapping a Redwood Forest with LIDAR
- Climate Change Threatens Aquarium Favorite
- To Save Animals, Put a Price on Them
- States Fight Pharmaceutical Industry Data Mining
- Google Launches Medical-Records App
- Crowdsourcing the Flu Vaccine
- Google.Org Spends $14 Million to Battle Pandemics

WiSci 2.0: Alexis Madrigal's Twitter , Google Reader feed, and project site, Inventing Green: the lost history of American clean tech; Wired Science on Facebook.