



## UvA-DARE (Digital Academic Repository)

### A data bias in interdisciplinary cooperation in the sciences: ecology in climate change research

Kwa, C.; Rector, R.

**Publication date**

2010

**Document Version**

Author accepted manuscript

**Published in**

Collaboration in the new life sciences

[Link to publication](#)

**Citation for published version (APA):**

Kwa, C., & Rector, R. (2010). A data bias in interdisciplinary cooperation in the sciences: ecology in climate change research. In J. N. Parker, N. Vermeulen, & B. Penders (Eds.), *Collaboration in the new life sciences* (pp. 161-176). Ashgate.

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

# A Data Bias in Interdisciplinary Cooperation in the Sciences: Ecology in Climate Change Research

Chunglin Kwa and René Rector  
University of Amsterdam  
c.l.kwa@uva.nl

Final Draft

Published in: *Collaboration in the New Life Sciences*, John Parker, Niki Vermeulen, Bart Penders (eds), London: Ashgate, 2010, pp. 161-176.

After 1980, and going full steam since 1990 or so, a number of sciences have reoriented themselves to data gathering on a scale immensely larger than in the post-World War II decades. We are witnessing a new ‘avalanche of data’, and fears of becoming ‘drowned in data’ have been repeatedly expressed during the past ten years (Conway, 2006).<sup>1</sup> The proximate cause of the data revolution is technology: the rapid development of medical imagery equipment, automated gene sequencing machines, and remote sensing equipment mounted on satellites orbiting the Earth are conspicuous examples. In response to the still evolving technologies of data collecting, new technologies of data handling, data storage, data retrieval etc., have come into existence. Databases have become instruments of scientific work, and their development and maintenance informs to an important extent ongoing research (Hine, 2006; Millerand & Bowker, 2009).

The new riches bestowed on the sciences in the form of data collecting and data handling technologies (many of them made available through public funding) point to an interesting paradox: the financial room for autonomous science at the universities is decreasing at the same time. The data-dominated sciences can often be located within large national and international programs. The International Human Genome Project, the Human Brain Project and many other such programs operate under criteria which do not come out of peer-reviewed ‘little science’ but are formed instead within state-funded bureaucracies and/ or entrepreneurial cultures. Science operates under a ‘new social contract’ with society (Jasanoff, 2005; Shapin, 2008). The terms of the new contract are to either lead to the development of marketable products or to contribute to more generally conceived societal goals, the elucidation of climate change and the mechanisms of global

---

<sup>1</sup> See Hacking (1990), for a description of what he termed the avalanche of data that hit the sciences in the early decades of the 19th Century.

warming obviously being a case of the latter. The big international programs may be said to fulfill a dual role: they help steer the academic sciences toward societal relevance while the management and the structures of the programs allow for transparency and accountability.

The more general point that we want to explore tentatively in this paper is whether a link can be established with the way sciences operate under the new social contract and their increased orientation toward data. During the period 1945-1980, the disciplinary organization of the sciences was linked to cultural prominence of high theory. Physics was at the top of the disciplinary hierarchy, and physics-envy invited the other sciences to mimic the deductive theory structure of physics. Today, the legitimacy of public funding of science is much less taken for granted than before 1980, inducing science policy officials to find new uses and audiences for what they once conceived of as autonomous or 'fundamental' science. They find, among other things, that at least some segments of the public identify 'sound science' with corroborated data rather than with theory.

We should not see, therefore, technological development in data collecting and data handling as a blind force. In the case of climate change research, NASA took the lead in defining the technological needs and standards for remote sensing equipment beyond the wildest dreams of climatologists, oceanographers and ecologists. For many practitioners of the latter sciences, the flood of geospatial data that was poured out over them seemed a mere instance of 'technology push'. But for NASA it was rational to engage in it, as it was for NSF and NOAA to ally themselves to NASA. Together, the three agencies forged the data revolution in the Earth sciences to an important measure (Kwa, 2005).

What we hope to elucidate in this chapter, however, is a more limited instance of the data revolution: its relation to new cooperative work patterns in science. We study a group of seven Dutch ecologists, who after 1990 all became involved with research on the influence of climate change on ecological systems, and we follow them to 2005. The primary vehicle that made them engage in climate research was an international scientific program: the International Geosphere-Biosphere Programme (IGBP), which in 1990 had become operational. As the ecologists involved amply testified to us, the IGBP led to a marked increase in their networking and planning activities. We observe furthermore that increased cooperation led to changes in their publication behaviour. After 1999, articles with more than ten authors with many different institutional affiliations in several different countries begin to appear on the Dutch ecologists' publication lists. In the articles, the use of databases is reported, recently constructed and much larger than previous databases constructed in field ecological research. We suggest that, at least in their case, the development of databases is an unplanned outcome of their participation in the IGBP. The cases of the Dutch plant ecologists are relatively modest in comparison to many much larger cooperative efforts and involving much larger databases. But that makes them no less interesting, as we will argue in this paper.

## 1 Global change on the ecologists' research agenda

The IGBP is a scientific program under the auspices of ICSU, launched in 1987 and currently planned until 2013. It groups notably the earth sciences, among which are included ecology, oceanography and geography, complementing the World Climate Research Programme (WCRP) (Kwa, 2005). The WCRP is dominated by meteorology. This program started in 1980 and it has a planning agenda that extends to 2015.

The planning of programs such as the IGBP takes place at many fora at the same time. Much of what is at stake during the planning process revolves around the question: which disciplines should be included, which disciplines should be left out? The WCRP did not want biological disciplines, therefore the IGBP would have them. Once the disciplines are in, they get shuffled around in subprograms of the big program, forming various interdisciplinary combinations, hopefully leading to interdisciplinary cooperation.

After the planning phase is over and the life of the program begins, the disciplines might become reshuffled. But once combinations are formed, room is made for bottom-up processes. Scientists at all levels of the IGBP are engaged in continuous programming efforts. The idea is that this will result in increased cooperation. The scaling-up of research is thought to follow from this process.

The primary internal organization of the IGBP is in Core Projects, one of which is the now terminated Global Change and Terrestrial Ecosystems (GCTE) (in which the seven Dutch ecologists participated). The IGBP-leadership selected for each of the Core Projects a leadership from among the world's top scientists. The Core projects further designated 'Foci', and again selected leaders for these, while the Foci were subdivided again into 'Activities'.

In addition, a number of cooperative networks were formed during the life of the GCTE, some formally under its umbrella, some not. The IGBP forms a world that is chock full of acronyms: NEWS (Network of Ecosystem Warming Studies), TERRACC (Terrestrial Ecosystem Response to Atmospheric Climatic Change, a network 'endorsed by the Global Land Project', the successor to the GCTE as Core project when the former was terminated in December 2004).

The IGBP itself has formed a partnership with three other international programmes: the WCRP, Diversitas, and the International Human Dimensions Programme, forming the Earth System Science Partnership (in the leadership of which is Rik Leemans, one of the seven Dutch ecologists central in this chapter). In addition, there are bilateral (or trilateral) contacts between the IGBP Core Projects, such as the BIG initiative of 1993, a cooperation between GCTE and LUCC (Land Use and Cover Change) and BAHC. End of 2003, GCTE was discontinued ('terminated') in order to merge with LUCC into the Global Land Project. The IGBP-leadership has not disclosed why the GCTE was terminated, but the reason was apparently to stimulate ecologists to further widen their

interdisciplinary outlook. The point of the coming and going of organizational platforms is to ensure that the process of programming, planning and evaluating is in continuous flux.

As a disciplinary group, ecologists have never been in the IGBP completely on their own terms. Even if the GCTE was originally built up according to the best disciplinary standards in ecology, it was clear that the program was to contribute to the big picture of climate change (Kwa, 2005). In one of the GCTE Foci, plants were exclusively considered as small pumps of carbondioxide.

Falling under ICSU, the IGBP is a ‘science’ organization. But the IGBP does not have money of its own. Instead, national research councils such as the US National Science Foundation and the Dutch NWO have committed themselves to funding IGBP research projects. In the absence of a governmental agreement on the IGBP, a bottom-up organization of research councils has come into existence: the International Group of Funding Agencies (IGFA) (Kwa, 2006). IGFA was founded in 1990, its aim is to ‘foster’ climate change research. To date, twelve countries participate in this informal science policy body. Member organizations subscribe to the activist agenda of IGFA with regard to climate change research. i.e. they acknowledge the importance of planning and programming. As a group, IGFA is a critical partner of the IGBP, keeping IGBP sharp by insisting on regular evaluation. Within their own countries, the respective funding agencies make it clear to individual scientists who apply for funding that they expect them to do their part in the programming activities on the various levels of the IGBP. Thus, NWO knows beforehand that a scientist’s project dealing with an aspect of climate change is part of a larger whole and embedded in an international network of researchers. For NWO, one of the founder members of IGFA, the special priority for the climate-related research coordinated by the IGBP was one of the very first occasions to substantiate its new strategic mission toward academic science, which it had received in 1986.

## **2 The effects of cooperation on seven Dutch plant ecologists**

We have limited our research to ecologists active in one of the IGBP core projects: the Global Change and Terrestrial Ecosystems Programme (GCTE). This entails a focus to plant ecological research, usually of a plant physiological nature and involving modelling. (A number of Dutch ecologists participated in other IGBP Core projects, notably Land Ocean Interactions in the Coastal Zone (LOICZ), in which we find among others fisheries researchers, and Land Use and Cover Change, which groups among others agricultural ecologists and geographers). Moreover, we have limited the number of ecologists whose work we report here to a group of seven. They consistently came out

first when in 2005 we requested a larger group of ecologists to name the scientifically most important and influential plant ecologists at Dutch universities and research institutes.

The seven plant ecologists have in common that they were already more or less established in research by 1990, when the IGBP was implemented internationally, and nationally shortly thereafter. At the time, some had already anticipated the coming importance of climate change, the others were soon to follow. They all could show the relevance of their work to the theme of climate change.

Jan Goudriaan and Siebe van de Geijn (Wageningen) had been investigating the physiological response of plant leaves to increased levels of carbon dioxide, and they also had a background in computer simulation modeling (they were physicists by training). The climate issue made them put the two together, by which they could help to address a vexing question at the time: the gap in the global carbon budget of the biosphere, also known as the 'missing sink': the atmosphere contains less carbon dioxide than might be expected on the basis of the increased fossil fuel combustion. Apparently the biosphere is sequestering the missing carbon, but where and how? This question had been an important reason to include ecology among the sciences of global change. Goudriaan was invited to become a member of the international Steering Committee of the GCTE from its inception.

Jelte Rozema (Vrije University) was likewise interested in basic aspects of carbon assimilation by plants. One group of plants metabolizes carbon differently than most plants and it turns out that this group will benefit comparatively from increased carbon dioxide in the atmosphere. This aspect of his research was soon to be found relevant to climate change, as was a general focus of his on arctic (Lapland Spitsbergen) ecosystems, an interest he shared with Rien Aerts (first at Utrecht University, after 2000 Vrije Universiteit), and Hans Cornelissen (Vrije University). Ad Huiskes (NIOZ, Yerseke) observed that in the Antarctic, temperature fluctuations appear as an important factor in affecting ecosystems sooner than in temperate regions, where there are simply many more factors of interest and the Antarctic is the most rapidly warming region in the world. In Huiskes' experience, the antarctic region "just happened" to be relevant to GCTE-themes, as if nature had applied relevant laboratory conditions there.

Aerts initially studied nitrogen deposition in subarctic peatbogs, but he found he could apply similar experimental and measuring techniques for studying carbon deposition, carbon having the more immediate salience for climate change studies. At his tundra research stations, he uses among other things small open-top chambers to increase mean air temperature by 1-2 degrees Celsius. Aerts made the transition from nitrogen to carbon in 1993. Cornelissen applied his interest in differences in functional plant traits to carbon storage by plants in the ground.

Rik Leemans was among the Dutch ecologists who by 1990 was already firmly established in climate research. An ecologist by training, he mastered the mathematical

techniques of simulation modeling later. But from then on his career was irreversibly wedded to the topic of climate change, more outspokenly so than the other ecologists discussed here. Leeman's first model was of a boreal forest, which soon was developed into a global model. He moved on to the RIVM in Bilthoven, which gave him a more 'applied science' institutional setting. Here he developed with others Image 2, one of the first influential 'integrated models' which study the impact of elevated temperatures on the world's ecosystems, including such aspects as forecasting changes of agricultural land use.

Reading through the publications of the seven ecologists, focussing on the publications in which they appear as single author or with one co-author, we conclude that their research interests in 2009 in terms of topics and the scientific questions they ask, are not fundamentally different from what they were doing before 1990. That is, they maintained a link to the disciplinary identity of ecology. With one exception: Leemans continued to broaden himself in interdisciplinary global change science. This observation does not imply in the least that the contributions to climate science by the other six were any less genuine than Leemans', only that in principle they could in the future contribute to other topics as well.

For both Goudriaan and Van de Geijn, the climate focus became too much of a straightjacket. Moreover, the scaling-up of research to a global level was hardly of interest to the rest of their respective research groups and Van de Geijn complained about the time spent on programming, themes to which we will return in the next section. By the end of the nineties, they withdrew from the GCTE.

Should we have interviewed younger scientists, whose Ph.D. work would have already been framed by Global change research, we may have had different findings to report. Younger scientists might have been socialized in global warming research to such a degree that changing fields would not have been an option. Apparently, for some of the older ecologists with disciplinary allegiances intact, global climatic change could in principle be seen as a transient subject. Yet, as we will argue in the next sections, in another sense the impact of the global change topic on ecology may be more lasting, through affecting ecology's style of research, and this might hold for both older and younger scientists. It would, therefore, be premature to conclude that the steering effect from the IGBP and NWO's support for it would have been transient. It is to the more lasting impacts of the 'aggregating' (Rip and Van der Meulen, 1996) effects of science policy that we will now turn.

All seven ecologists indicate that membership of a GCTE-related forum was necessary toward NWO because it functioned as an indication of one's scientific quality. Apparently, a similar situation existed in other IGFA-countries. There was at any rate uneasiness in the higher echelons of the GCTE, where it was felt that GCTE-

membership was taking over the function of ordinary peer review.<sup>2</sup> But even when the GCTE-leadership expressed the opinion that this should not be the case, adoption of the research projects by the GCTE organization continued to work as equivalent to bearing the stamp of scientific approval. As some of our respondents noted, this worked not only for NWO but on the EU-level as well. At any rate, a decision not to take part in the GCTE was perceived by the ecologists as harmful to one's interests.

For one Dutch climate research program, not administered by NWO, IGBP-membership was a formal requirement to be eligible for funding. The members of its panel, among whom were several of the seven ecologists, found themselves in the awkward situation that applying this condition implied granting all the available funds to themselves, as they were the only scientists meeting the requirement.

Nearly all our respondents complained about the time involved in participating in IGBP committees, with the exception of Leemans, for whom the changing organizational landscape of the IGBP provided an opportunity of continued learning about new aspects of climate change. But, as noted, for Goudriaan and Van de Geijn the time-consuming planning activities were additional reason to withdraw.

Yet participating in the planning activities of the IGBP was not merely a time consuming form of window dressing. The planning activities did not fail to have aggregation effects at the national levels (with the Research Councils as the most important patrons of the IGBP), and on various international levels as well, such as the EU, in which typically four groups would participate in a program funded through the 6th and 7th Framework Programmes. Rozema participated with partners from Slovenia, Sweden and Germany in the UVAQTER project.

Increased cooperation can also be read from the lists of publications of several from among the seven ecologists. With Rien Aerts, for instance, publications by five authors and more begin to appear in 1993, followed by publications with ten and more authors after 1999. A similar pattern can be seen with Jelte Rozema and Cornelissen. The publications with ten or more researchers are especially interesting because the authors are from a variety of institutions in a variety of countries. It is these publications that most significantly are organized around shared databases, the theme of the next section.

However, publications of ten authors and more do not feature on the publication lists of Goudriaan and Van de Geijn. This may come as no surprise, given their views on the drawbacks of international cooperation. They do appear as authors of articles with six researchers or so, but in these cases, their fellow authors are all from Wageningen.

### **3 The orientation toward data**

---

<sup>2</sup> Ian Woodward, letter to Brian Walker, 1 december 1993. (Personal archives of Jan Goudriaan).

We now come to the central claim of our paper, which is that the science policy initiatives toward the sciences help foster the data revolution in the sciences. This is not a stated aim of science policy, and neither could it be. The IGBP as a ‘science’ organization, and IGFA, representing ‘policy’, takes great care in upholding the boundaries between science and policy. If in practice the boundaries between science and policy are transgressed fairly routinely (see Kwa, 2006), their formal separation is respected. Science policy may promote certain priority areas for science, such as global change, but the way science is conducted and its internal quality control is an affair of science only. Yet at the same time the entire planning phase of the global change research programs had been framed by the new possibilities concerning data collecting technologies through remote sensing.

Looking at the early history of the IGBP, it looks as if the first impulse to the data-orientation of the program had come from scientists. At a planning workshop held at NCAR, in the US, it had been stated: ‘Among the technological drivers are advances in the remote sensing of the Earth by spacecraft, (...) and the capacities of modern computers and communication networks to store and disseminate vast quantities of data and, through common access to these stores, to open new avenues of interdisciplinary exchange’ (Atmosphere/Biosphere Workshop, 1985, p. 1). The following year it was said that ‘Remote sensing observations of land and ocean areas from space hold the greatest hope for observing significant global changes... Observations of the Earth from space are an essential element of the IGBP, and the reason, as much as any other, why the programme is now possible’. (Ad Hoc Planning Group 1986, pp. 7-8).

Who made these statements? Speaking here are elder statesmen of science, mostly meteorologists, and a few ecologists, reacting to an opportunity offered by NASA to make use of their satellite technology for a ‘mission to Planet Earth’. But not only did they recognize the new technology’s potential for data generation, they also saw the opportunity which data provided for interdisciplinary integration, through making scientists across disciplines work with the same set of data.

As we have described elsewhere (Kwa, 2005), three science policy individuals, Robert Corell at the NSF, and two more at NOAA and NASA respectively, shaped the US Global Change Research Programme. The GCRP houses the American contribution to the IGBP, which, as initially seen from the US perspective, was as much the international umbrella to the GCRP as its international extension.

Corell was an early advocate of good data management within the GCRP and the IGBP. He contributed to the establishing of data facilities for the various climate programmes. The IGBP installed a Data Information System (IGBP-DIS) on a par with the Core Projects. (It has now been transferred to the ESSP).

For the science policy individuals which Corell would unite around himself in IGFA, the idea of interdisciplinary integration through a shared use of geospatial data was an attractive one. It would help justify the expensive equipment put up in outer space, and it would ensure the scaling-up of ecological research to meet the requirements of

meteorologists and other earth scientists who were already working on a global scale. In a word: the shared use of data would bring focus and critical mass in global change research, avoiding the cumbersome and unpractical route of theoretical integration across the disciplines.

Several US ecologists organized themselves as a 'user group' toward NASA, specifying requirements for remote sensing instruments. The group interacted with NASA during several years, and the process was facilitated by ecologists on the NASA staff (Van Rees, 2007). They reported on their efforts in *Ecology*, the scientific journal of the Ecological Society of America (Matson, 1991; Roughgarden, Running & Matson, 1991). In France, an ecologist was entrusted with the leadership of a spatial research laboratory, replacing a physicist. Evidently, he could leave his mark on the development of remote sensing equipment in the French space program.

The Netherlands does not have a space program of itself, but it participates in the European Space Agency. The Dutch government supported the use of remote sensing technology by the Dutch scientific community (NRSP, 1986; NRSP, 1990). One of the science policy officials who had been in charge with the preparation of the two consecutive programmes (1986-1990 and 1990-2000) was John Marks, an early advocate of the IGBP and organizer of IGFA, along with his US counterpart Robert Corell. Marks did not attempt to direct Dutch IGBP scientists to the use of remote sensing (Marks, interview with author). The first National Remote Sensing Programme made it clear that it was motivated in part to enhance Dutch scientific quality in the fields of meteorology, oceanography and climate research (NRSP, 1986, p. 15). Dutch geographers and ecologists were not slow in catching up with the possibilities of remote sensing technologies. Between 1990 and 2000, for instance in the field of riverine studies, remote sensing had developed 'from an additional source of data to an indispensable one' (Van Hemert, 2008, p. 46).

In the case of geospatial data, we may therefore conclude that there is an intrinsic link between the data revolution in the earth sciences and the science policy context of the large international scientific programmes. But the data problematique of the seven Dutch ecologists lies elsewhere. Even when they are (occasional or regular) users of geospatial data, they are primarily concerned with the field data they generate themselves. During the period 1990-2005, there is a gradual increase in the importance assigned to field data as a distinct area of concern, most visible in references to larger databases they share with co-authors of their articles.

Until fairly recently, databases containing field data were the personal property of a vast array of individual principal investigators. Starting toward the mid-1990s, however, field data and databases constructed on the basis of them became the subject of policy. Grant giving agencies now request the publication of databases and organizations such as the Ecological Society of America have formulated criteria for published data to enable their use by others in other contexts (Mitchener et al., 1997; Millerand & Bowker, 2009). How

and why have these policies come into being? The story is different for different countries. It is worth noting that that in the US case, the NSF is heavily informing field database policies, through their Long Term Ecological Research Program (LTER). It is equally worth noting, however, that no institutional ties exist between the US GCRP and the LTER program. Field database policies originated in a learning-by-doing fashion, to some extent as the unintended outcome of increased cooperation between the field scientists entailed by international and national programs. If this was true for US ecologists involved in the LTER sites, it certainly applied to the seven Dutch ecologists.

#### **4 Databases come into two different kinds**

Within the group of Dutch ecologists, Rik Leemans is a special case, as he assembled data from existing data. He built a large data-base in cooperation with Wolfgang Cramer, in the eighties, at the International Institute for Applied Systems Analysis (IIASA), near Vienna. Initially, they reworked meteorological data to suit their model of boreal forests, but soon found they could extend the database with other regions of the world as well. The enlarged database was put to use in a model which simulated the impact of climate change on terrestrial ecosystems. Leemans moved on to the RIVM in Bilthoven, the Netherlands, where he developed IMAGE 2, also a climate impact model, which would acquire a very visible policy relevance. (It was used during the negotiations at Kyoto, leading to the Kyoto protocol).

Van de Geijn and Goudriaan generate their own datasets in the experimental settings they have developed. Particularly noteworthy is the Wageningen Rhizolab, a large research facility for the study of soil-root-shoot-atmosphere relationships in crops. The Rhizolab is equipped with 160 sensors which measure simultaneously a number of variables, fully automated. A data acquisition control unit automatically collects the readings from the sensors. They note that “like in most automatic data acquisition systems, the constant stream of data generated can easily exceed the investigators ability to process it” (Van de Geijn et al, 1994, p. 286). On a ‘micro’-scale, they face problems quite similar to those of researchers who work with remote sensing data.

The remaining four Dutch ecologists who are central in this article also collect their own field data. Their style of research is experimental, too, hence the data Aerts, Cornelissen and Rozema gather on plant traits are collected on the spot at their respective experimental stations in sub-polar regions, and/ or in their carbon enrichment open top chambers in the Netherlands. Their most recent papers contain suggestions that they might use data obtained through remote sensing in the near future, but to our knowledge, they do not yet at this point. But the evolution of their research over the past ten years shows an increasing emphasis on the development and deployment of databases of field

data. This trend is especially visible in the multi-author articles, notably those with ten or more authors.

By 2005, at the time of our interviews, Rien Aerts was solidly aware of the pressure on individual ecologists such as himself to the generalized use of data from his experiments. Hence, his membership of the International Tundra Experiment (ITEX), which prescribes standardized protocols for experimental procedures of the tundra researchers. ITEX groups scientists from 11 countries.

Hans Cornelissen contributes to Glopnet, a database maintained at Macquarie University, Australia, comprising plant trait data from 175 sites around the world (Wright et al, 2004). More recently, Glopnet data have served as input to Dynamic Global Vegetation Models. Glopnet is a bottom-up organization, it ‘encourages’ scientists to make their data available, but has no means to enforce it. Because of the highly complex nature of Glopnet data, there is as yet no golden metadata standard and internal discussions about data policy continue.

## **5 The internal taxonomic logic of databases**

In experimental science, regardless of its specific subject, each experiment testing a hypothesis typically generates its own data. In principle, this is true as much for physics as it is for the experimental botany of Aerts, Rozema or Van de Geijn and Goudriaan. Yet ecologists such as Aerts and Rozema face a charge which is typical for ecology: are the observations or outcomes of the experiments generalizable beyond the local features of the ecosystem which they happen to study? (see also Kohler, 2002). The larger context of climate change research makes this question obvious. The scaling up of the investigations is a requirement from which no ecologist can escape.

But we should note that, in principle, two sorts of generalizations can be made: one is from the (published) case studies, the other is directly from the data assembled for the case studies. The data revolution invites to the second type of generalization. However, the reuse of data by other researchers, in some cases may be decades after they were assembled, places high demands on the format of the data. As Mitchener (1997) notes, there is a lot of tacit knowledge around field data that needs to be made explicit. For instance, data are the product of measurement apparatus different from apparatus used by other researchers, hence the data should be calibrated and standardized. Therefore, data about data (“metadata”) should be provided.

By the time the problem of standardization of field had been identified, the data problem had also become apparent for geospatial data. In 1993, the absence of an agreed-

upon format of remote sensing (geospatial) data for ecology was noted with astonishment by the chairman of the IGBP-DIS standing committee.<sup>3</sup>

Making data accessible to other researchers requires a degree of standardization and organization which is not usually achieved by individual researchers alone. For geospatial data it has been noted that producers and users are in the process of becoming separate communities (Conway, 2005).

A transition from an experimental style, embedded in hypothetical-deductive arguments, to a more data-oriented style has been noted by several authors (Balmer, 1996; Beaulieu, 2001). A science dominated by data requires methodologies and practices different from the better known experimental sciences: questions of classification come to the fore (Bowker & Star, 1999; Bowker, 2005; Hine, 2008).

The field of genomics is especially interesting. Since Craig Venter developed the ‘shot-gun’ technique to sequence genomes, a data-driven technology which replaced the earlier hypothesis-cum-experiment approach of investigating genes one by one, billions of data still wait for analysis and interpretation (Balmer, 1996; Leonelli, 2007).

In her analysis of the *Arabidopsis* research community, Leonelli (2007) noted an interesting split. Of central importance to this field of plant genomics is a database, to which researchers worldwide contribute, and which they can access for their own research questions. The point of such a database is that data can travel, from the contexts in which they were generated to new unforeseen contexts of application. While in the original experimental context, experimental conditions and specific theoretical expectations determined which data were relevant and which were not, no such backgrounds can be drawn on within the database. Instead, new requirements need to be met: the ability of data to be retrieved from a user-friendly interface. This point becomes salient when curators are appointed to administer the database. Curators make their decisions concerning which data to include in the database and which not, which labels to attach to the data and which not. We may call the ‘logic’ of data handling and storage taxonomic, rather than hypothesis-driven (Kwa, 2010, forthcoming).

Is a similar process underway with regard to ecological field data? At least the LTER network has made steps in this direction by appointing a Network data manager at its Central Office (since 1989). Also in 1989, the first issue of an electronic magazine ‘Databits’ was published, to enhance cooperation between the Network Data Managers of the respective sites. While these individuals probably define themselves as service personnel with regard to the field ecologists, and not (yet) as a distinct research community of data curators, the formulation of ‘metadata’ policies might be a decisive step toward the development of a taxonomic research style. At present, it is still a contested issue. It has been noted time and again that many ecologists are reluctant to share their data, to say the least. The review committee of the “20-year LTER Review”

---

<sup>3</sup> John Townsend, fax to Brian Walker, 29 July 1993.

wrote that “the first and fundamental strategy must be the organization of LTER research *a priori* by hypotheses and theory, with networked data acquisition, analysis and testing by predictive models across broader and broader phenomena.” Theory first, data second. Yet the Committee recommended that LTER should establish informatics as a core function to integrate LTER data and that the NSF should increase the importance of data management and informatics in its evaluations of the LTER (NSF, 2002). At present, NWO, the Dutch counterpart of NSF, seems somewhat less proactive toward Dutch ecologists with respect to data, but the difference is in degree only. It is conceivable that we will have a new ‘taxonomic’ ecology in the foreseeable future, taxonomic because it relies on the cataloguing of data. There would have been a science policy influence on three levels. First: by enabling the data revolution through the promotion of the development of data collecting and data handling technologies. Second: by directly speaking to scientists on the importance of data. Third: by creating the conditions for large-scale cooperation between scientists, with data as the primary means of achieving integration.

## List of acronyms

BAHC	Biospheric Aspects of the Hydrological Cycle
ESSP	Earth System Science Partnership
GCRP	Global Change Research Program (US)
GCTE	Global Change and Terrestrial Ecosystems (IGBP)
ICSU	International Science Organization (formerly International Council of Scientific Unions)
IGBP	International Geosphere-Biosphere Programme
IGBP-DIS	IGBP - Data and Information Services
IGFA	International Group of Funding Agencies
IHDP	International Human Dimensions Program
LOICZ	Land-Ocean Interactions in the Coastal Zone (IGBP)
LTER	Long Term Ecological Research Sites (US)
LUCC	Land Use and Cover Change (IGBP)
NASA	National Aeronautics and Space Administration (US)
NIOZ	Netherlands Institute for Sea Research
NRSP	National Remote Sensing Programme (Dutch)
NOAA	National Oceanic and Atmospheric Administration (US)
NSF	National Science Foundation (US)
NWO	Netherlands Organization for Scientific Research
RIVM	National Institute for Health and Environment (Dutch)
UVAQTER	UV Radiation in Aquatic and Terrestrial Ecosystems
WCRP	World Climate Research Programme

## Acknowledgements

We thank Jan Goudriaan for putting his personal IGBP archive at our disposal, and two reviewers for their constructive criticism.

## References

- Balmer, Brian (1996), 'Managing Mapping in the Human Genome Project', *Social Studies of Science* 26, 531-573.
- Bowker, Geoffrey C. (2005), *Memory Practices in the Sciences*, The MIT Press, 2005.
- Bowker, Geoffrey C. and Susan Leigh Star (1999), *Sorting Things Out: Classification and its Consequences*, The MIT Press, 1999.
- Conway, Erik M. (2006), 'Drowning in Data: Satellite Oceanography and Information Overload in the Earth Sciences', *Historical Studies in the Physical and Biological Sciences* 37, 127-151.
- Hacking, Ian (1990), *The Taming of Chance*, Cambridge: Cambridge University Press.
- Hine, Christine (2006), 'Databases as Scientific Instruments and their Role in the Ordering of Scientific Work', *Social Studies of Science* 36 (2006), 269-298.
- Jasanoff, Sheila (2005), *Designs on Nature*, Princeton, NJ.: Princeton University Press.
- Kohler, Robert F (2002), *Landscapes and Labs: Exploring the Lab-Field Border in Biology*, Chicago: The University of Chicago Press.
- Kwa, Chunglin (2005), Local Ecologies, Global Science: Discourses and Strategies of the International Geosphere-Biosphere Programme, *Social Studies of Science* 35, 923-950.
- Kwa, Chunglin (2006), Interdisciplinary Research through Informal Science-Policy Interactions, *Science & Public Policy* 10, 457-467.
- Kwa, Chunglin (2011), *Styles of Science: A History of Science from Antiquity to the Present*, Pittsburgh: University of Pittsburgh Press.
- Florence Millerand and Geoffrey C. Bowker (2009), 'Metadata Standards: Trajectories and Enactment in the Life of an Ontology', in: M. Lampland & S. L. Star (Eds.), *Standards and their Stories*, Cornell University Press, 2009.
- Leonelli, Sabina (2007), 'Weed for thought: Using *Arabidopsis thaliana* to understand plant biology', PhD dissertation, Vrije University, Amsterdam.
- Matson, Pamela (1991), 'The Future of Remote Sensing in Ecological Studies', *Ecology* 72, 1917.
- Mitchener, William K. et al. (1997), 'Nongeospatial Metadata for the Ecological Sciences', *Ecological Applications* 7, 330-342.
- NRSP (1986), *Nationaal Remote Sensing Programme*, Tweede Kamer der Staten Generaal, The Hague.

- NRSP (1990), *Nationaal Remote Sensing Programme 1990-2000*, Tweede Kamer der Staten  
 Generaal, The Hague.
- NSF (2002), LTER Report.
- Rip, Arie & Barend van der Meulen (1996), 'The post-modern research system', *Science and  
 Public Policy* 23: 343-352.
- Roughgarden, J, S.W. Running and Pamela A. Matson (1991), 'What does Remote Sensing  
 Do for Ecology?', *Ecology* 72, 1918-1922.
- Shapin, Steven (2008), *The Scientific Life: A Moral History of a Late-Modern Vocation*, Chicago:  
 Chicago University Press.
- Van de Geijn et al (1994), The Wageningen Rhizolab - a facility to study soil-root-shoot-  
 atmosphere interactions in crops, *Plant and Soil* 161, 275-287.
- Van Hemert, M (2008), 'Making Rivers Modular: Emerging River Science 1980-2005',  
 PhD dissertation, University of Twente.
- Van Rees, J (2007), "The Earth Observing System and Ecology: Technology setting the  
 stage for a changing discipline", unpublished MSc Thesis, University of  
 Amsterdam.
- Wright, Ian, and Reich P.B., Westoby M., Ackerly D.D., Baruch Z., Bongers F.,  
 Cavender-Bares J., Chapin F.S., Cornelissen J.H.C., Diemer M., Flexas J., Garnier  
 E., Groom P.K., Gulias J., Hikosaka K., Lamont B.B., Lee T., Lee W., Lusk C.,  
 Midgley J.J., Navas M-L., Niinemets Ü., Oleksyn J., Osada N., Poorter H., Poot P.,  
 Prior L., Pyankov V.I., Roumet C., Thomas S.C., Tjoelker M.G., Veneklaas E.,  
 Villar R., The World-Wide Leaf Economics Spectrum, *Nature* 428 (22 April 2004),  
 821-827.