Australian Plant Phenomics Facility

NEW TOOLS AND A NEW PARADIGM FOR RESEARCH

While much scientific research was once limited by lack of data due to cost or technical complexities in its acquisition, in the past 15 years many of these obstacles have fallen and will continue to do so. The fact that we can now measure the world in high resolution through time enables a fundamental shift in how we can undertake science and research.

However, data interoperability, data discovery, and tools for interacting and analysing complex data all present new and growing challenges, and solving these challenges in a wide-scale, cross-domain fashion is likely to be essential to effectively tackling the grand challenges facing Australia in this century.

Meanwhile, we still largely lack the tools for easily working with data that is multi-layer, high complexity, high dimensionality, and high time and spatial resolution. In short, Australia's research community needs the equivalent of an 'Office suite' of tools that would let us make best use of these data, and the training to confidently embrace using them.

Film and gaming technologies for real world modelling for research

In the past few decades, the rapid pace of change in computational power, computer vision and 3D modelling have enabled the film and game industries to develop extremely powerful tools for building realistic 3D environments. This has created new opportunities for scientific modelling that have not yet been well exploited by the research community. Open world video games, for example, where the user is free to explore an entire environment, typically encompass hundreds or even thousands of kilometres of virtual landscapes (or in the case of some space exploration games, billions of planets all with their own unique ecosystems and life forms)[i].

Since it would be impossible to place every tree and animal by hand in landscapes of this scale, a technique called procedural modelling is used. In procedural modelling, the computer is given a recipe for automatically generating a visually realistic landscape. As games have become more complex and large scale, procedural modelling techniques have developed to the extent that huge land areas can be generated with realistic geology and animal and plant life that follow basic biological principles (for example mixed species forest grow up based on interspecies competition or tall grass growing at the edge of a stream will grow towards the areas of highest light). This is a very low overhead way to generate complex landscapes.

The game holds a set of 3D models or 'assets' of the objects in an environment, such as different species of plants and trees. The models are designed in such a way that they can be varied and grown to different sizes dynamically using growth models provided by the developer. Thus, rather than having to store a full model of the entire landscape on disk, the program just stores an outline of the landscape with a list of the various biomes and creatures present, along with 3D models of all the plant and animal species and parameters associated with that area including plant growth and animal behaviour patterns. When a user enters a new area in the game, the game 'grows' the area in an 'on-demand/real time' way, from the seed data provided.

While these tools are now used extensively by the game and film industries, taken in a research context these tools provide a number of unique opportunities to the science community that have not been previously available. For example, automated workflows already exist to import a 3D digital elevation model (DEM) of a landscape and populate it with plants and trees. Rather than using the 'cartoon' growth models as deployed by the gaming industry, for scientific contexts it is possible to use biologically robust models for growth and competition and to integrate impacts of factors such as aspect, soil composition, and water and light availability.

i. For example, the open universe simulation game No Man's Sky encompasses a universe of more than 18x10^18 unique planets with unique ecosystems and life forms that are procedurally generated deterministically, meaning that any player going any of the 18 quintillion planets will encounter the same planet and life forms as anyone else. See the article: https://en.wikipedia.org/wiki/No_Man%27s_Sky#Gameplay

Screenshot of a procedurally generated forest from the Ubisoft game Far Cry 5.

Potential of real-world modelling of Australia's landscape

Given that high-resolution, high-frequency time-series data are now available, characterising many facets of Australia, it would be possible to build a map of the entire continent where every farm, forest and field was generated procedurally, on the fly, when the user requests it. The procedural generation of the natural environment could be based on combining satellite and drone data with scientifically robust growth models and expertly-informed and AI-supported tools for interpreting them.

For a practical example of how this would be useful, imagine that we want to build tools to help agricultural production in Australia shift to being net carbon negative. It is now possible to build an app where the user puts in the address of their farm, and the digital elevation model (DEM) of the farm is downloaded from public data. Roads, vegetation and water bodies can be added to the model using public data sources and machine learning. Estimates of soil moisture, soil and plant health can be added using satellite data, machine learning techniques, and if the data aren't adequate, by on the ground sampling.

The physical landscape, plants and even animals can be modelled in 3D, then automated scenarios can be modelled to recommend improvements. For example, the system might identify how new contouring can be done to improve water retention on the landscape or suggest which application of regenerative agricultural methods might yield highest returns. The models of the landscape can be run casting forwards in time under various climate models to predict how different management practices might affect income and long-term carbon storage for example. With robust open access modelling and a carbon credit system in place, it could potentially be possible for banks to create apps to let farmers determine if they qualify for green loans to improve their property while increasing carbon drawdown at scale across Australia.

Virtual and Augmented Reality

While film and gaming software may help us effectively model new complex datasets in 3D, we can hardly use our conventional software interfaces for doing the sort of 'Next Gen' science that these types of data allow. The 3D visualisation capabilities offered by Virtual, Augmented and Extended Reality technologies (VR, AR, XR)[ii] can enable the next generation of tools for working in situ with living, changing 3D models of the real world and other complex data sets.

Humans have evolved for millennia as visual creatures and we use sight as a primary sense for finding meaning.

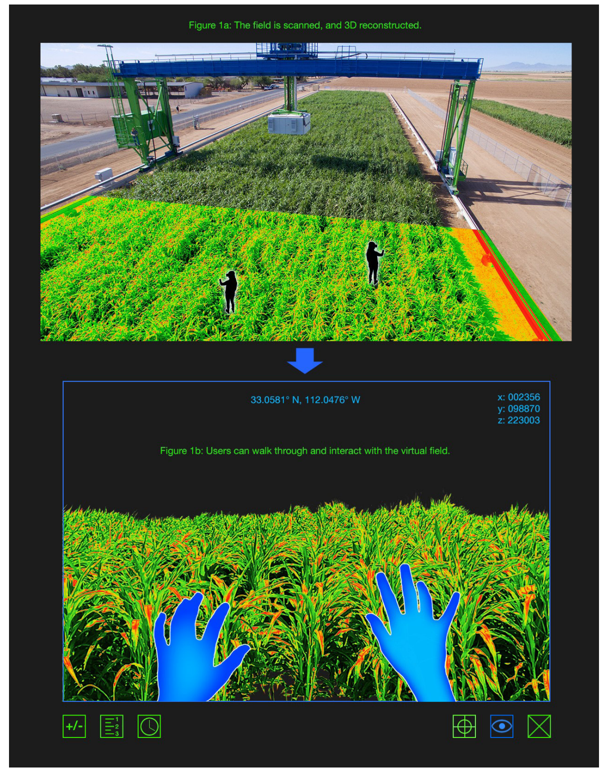
A key aspect of the value of research field work is not the measurements one makes, but the time spent immersed in an ecosystem, and the intuition it provides, that informs the direction of our data collection and quantitative analysis. It is an open secret in research that major breakthroughs don't usually come from looking at a particular figure or graph, instead it is often only when one pairs the quantitative measurements with the big-picture intuition gleaned from hundreds of hours spent immersing oneself in thinking about a particular research problem, that new insights emerge.

ii. In Virtual Reality, the user dons a pair of goggles and is completely immersed in a virtual world. With Augmented Reality, the user wears a headset that is more like a pair of glasses and so the user sees the normal world through the AR with 3D holographic objects projected into it (e.g., a 3D plant or farm model on the users' desk). Some VR headsets can also show the user a version of the real-world using cameras embedded in the VR goggles, and then virtual objects are added into the space. XR is used as a catch-all term to capture all of the technologies and potential ways that people might combine virtual and augmented realities.

Part of what the XR approach enables is letting researchers use their visual intuition to see if what was being measured actually reproduces the phenomena they are seeking to understand. Humans naturally use their intuition to extrapolate quite complex hypotheses from sometimes very low-resolution data. This can be both advantageous and problematic [1, 2], and probably happens to a much greater degree in research than people realise. Consequently, we need tools that let us use the data we have to reproduce a digital version of what we are measuring and then to perform the same sorts of data collection and analyses we do in the real world to see how well our data supports what we imagine ourselves to have measured. An additional utility of this approach is that integrating all our data into a single model requires the data to be interoperable in ways that in practice, it frequently is not.

Once we have these sorts of dynamic 3D models of our data, XR tools will let us interact with them in a 'hands-on' intuitive manner. For example, when a crop researcher is pulling up the data about a particular field trial, they would do so by putting on lightweight AR glasses and opening an actual 3D model of the site with all the data overlaid. They could view the entire research site as a model on their desk or zoom into a particular area of interest. If they wanted to be more immersed in what took place at that location, they could expand the model into an immersive 1:1 scale model and walk around inside the data.

This approach doesn't imply that conventional analytical tools are no longer useful, instead this approach augments traditional analysis with additional tools that are better suited for working with complex, dynamic 3D datasets. For example, a plant scientist might be working on their computer, analysing a numeric dataset capturing rate of 3D growth in their study plant in a growth chamber, relative to a range of biotic and abiotic factors. They may find a region of the data that isn't behaving as expected and so they open up the full 3D dataset of the plants onto their desk, to better explore visually what was actually happening. In addition to a regular colour model of the plant, they could change the surface of the plant to show information on photosynthetic activity or thermal layers, or un-randomise an experiment to group all plants of a particular genotype together with associated metadata for each group displayed above it. This way they can easily pair conventional analyses with newer methods of viewing complex 3D time-series data. A further benefit of this approach is that users can easily interact with remote collaborators. So, for example a researcher at a university in Australia could enter a virtual space with a collaborator in the US and they could explore the outcome of a set of experiments together. These types of applications are particularly intriguing in a potential future where international travel is reduced. The enhanced data visualisation opportunities offered by XR also provide significant education and outreach opportunities both in terms of sharing existing research and in creating virtual research or lab experiences that are significantly more immersive than traditional computer based 'virtual' labs.



VR/AR interface for enabling NextGen analysis of agricultural data and genomics. Image created by William Kezele, Danforth Plant Sciences for Tim Brown, APPF.

Digital twins of our real world

If we consider the first revolution in digital connectivity to be the Internet and the World Wide Web, and the second one to be digitisation of our social graph and social network, then a third and current revolution we are now going through is the digitisation of reality as a whole (the mapping and indexing of the entire world in 3D).

Although well-designed digital maps have been around since Google Maps launched 15 years ago, this new trend is not about 'mapping' in the traditional sense of helping people get from one location to another, but 'mapping' in the sense of creating a digital 3D copy of the entire world with every real physical element (mailbox, tree, boulder, house, etc) having a so-called 'digital twin' in the cloud. This approach is already being widely adopted both at the city and national level to improve city and disaster planning [3,4] and even internationally to empower citizens, governments and industries across multiple countries to collectively share the responsibility to monitor, preserve and enhance marine and coastal habitats [5]. For example, as part of the 'Open Decade Call For Actions' program, the EU is currently investing millions of euros to create a *"Digital Twin of the Ocean integrating all European assets related to seas and oceans (data, models, physical ocean observatories at sea) with digital technologies (cloud, High Performance Computing [HPC] capacities, Artificial Intelligence [AI] and data analytics) into a digital component that represents a consistent high-resolution, multi-dimensional and (nearly) real-time description of the ocean"* [6]

An impressive Australian example is the recently released Digital Twin of Adelaide which is a fully rendered, spatially accurate model of $1000km^2$ including trees and vegetation [7]. (While the people and cars in the Adelaide model are simulated for visual effect, it should be noted that in addition to the physical map of Adelaide, nearly everyone with a smartphone is currently being tracked to sub 5m resolution by at least their mobile company, and every state government knows what nearly everyone looks like as well as what vehicles they own. It wouldn't be a stretch to populate this model with near real-time models of nearly everyone in Adelaide, including their make and model of car or when they are travelling by mass transit or public scooter etc.)

For this paper, the term Digital Twin captures the concept of building persistent maps of the world that, in addition to 3D models of cities containing static data, provide the capacity to track and index everything taking place at any site, through time, wherever this data is being recorded, as well as tools for automating the entry of new data and accessing all the data programmatically.

The range of applications enabled by these sorts of maps is immense for governments, organisations and individuals, particularly when coupled with AR applications on mobile devices and AR wearables (with these devices expected to become widespread in the next decade). Powerful applications include 'intelligent exploration' of possible futures, such as city planning use cases where one might virtually insert a proposed building into an existing map of a city to understand its impact on views and ambient light in adjacent buildings, or the modelling impacts of proposed construction on human movements via modelling crowds or traffic within the digital twin of a city. In the environmental domain, one can imagine combining satellite and on the ground data from Australian sites to build continent-wide realistic 3D models of the 2019/2020 bushfires to enhance future bushfire response planning or as a visualisation tool to help the public better understand how drought and climate contribute to bushfire impacts.

References

1. Marie Tharp's ground-breaking maps brought the seafloor to the world https://www.sciencenews.org/article/marie-tharp-maps-plate-tectonics-seafloor-cartography

2. Can We Rely on Our Intuition? https://www.scientificamerican.com/article/can-we-rely-on-our-intuition/

3. Telecommunications and utilities infrastructure in the Spatial Digital Twin https://www.spatial.nsw.gov.au/what_we_do/projects/digital_twin

4. Data61 Building Australia's National Digital Twin with Cesium https://cesium.com/blog/2019/07/17/data61-building-australia-national-digital-twin/

5. A transparent & accessible ocean: Towards a Digital Twin of the Ocean https://ec.europa.eu/info/sites/info/files/research_and_innovation/green_deal/gdc_stakeholder_engagement_topic_09-3_digital_ocean.pdf

6. Transparent & Accessible Seas and Oceans: Towards a Digital Twin of the Ocean https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/lc-gd-9-3-2020

7. A virtual clone of Adelaide for digital twin and smart city applications https://www.unrealengine.com/en-US/spotlights/a-virtual-clone-of-adelaide-for-digital-twin-and-smart-city-applications