

Cultivating Responsible Plant Breeding Strategies: Conceptual and Normative Commitments in Data-Intensive Agriculture



Hugh F. Williamson and Sabina Leonelli

Abstract This chapter argues for the importance of considering conceptual and normative commitments when addressing questions of responsible practice in data-intensive agricultural research and development. We consider genetic gain-focused plant breeding strategies that envision a data-intensive mode of breeding in which genomic, environmental and socio-economic data are mobilised for rapid crop variety development. Focusing on socio-economic data linkage, we examine methods of product profiling and how they accommodate gendered dimensions of breeding in the field. Through a comparison with participatory breeding methods, we argue that the conceptual commitments underpinning current methods of integrating socioeconomic data into calculations of genetic gain can preclude the achievement of key social development goals, and that better engagement with participatory approaches can help address this problem. We conclude by identifying three key avenues towards a data-intensive approach to plant breeding that utilises the diverse sources of relevant evidence available, including socio-economic data, and maximises the chance of developing sustainable and responsible strategies and research practices in this domain: (1) reliable, long-term management of data infrastructures; (2) ongoing critical analysis of the conceptual foundations of specific strategies; and (3) regular transdisciplinary consultations including expertise in the social studies of agricultural science as well as participatory breeding techniques.

Keywords Plant breeding · Genetic gain · Responsible research practice · Participatory methodologies · Gender · Precision agriculture

H. F. Williamson (✉) · S. Leonelli

Exeter Centre for the Study of the Life Sciences (Egenis), University of Exeter, Exeter, UK
e-mail: h.williamson@exeter.ac.uk; s.leonelli@exeter.ac.uk

© The Author(s) 2023

H. F. Williamson, S. Leonelli (eds.), *Towards Responsible Plant Data Linkage: Data Challenges for Agricultural Research and Development*,
https://doi.org/10.1007/978-3-031-13276-6_16

301

1 Introduction: Data-Intensive Breeding for Accelerated Genetic Gain

As the previous chapters in this volume have made abundantly clear, the impact of data-intensive tools and methods on plant science and agriculture is extensive and motivated by a wide variety of goals ranging from reducing human labour to tracking dangerous pathogens, increasing yield, identifying agronomically promising plant varieties and understanding the impact of environmental and climatic changes on cultivation and food systems. Here we aim to reflect on some of the conceptual assumptions underpinning the implementation of data-intensive technologies and genetic insights in the agricultural domain. Specifically, we focus on one emerging trend for how plant breeding systems could be built to combine big data availability, including genomic, environmental and socio-economic data, with models for genomic prediction and specific selection methods. This trend is organised around the widespread adoption of genetic gain as a key indicator for evaluating and monitoring the outcomes of plant breeding, and for designing plant breeding strategies and seed system interventions for the future. In our view, this includes specific conceptual and normative commitments to a particular vision of agricultural development, which need to be explicitly drawn out to ensure that the strategies used to realise such a vision within specific situations are both scientifically reliable and socially responsible.

The rate of genetic gain is a statistical measure of the change in a population average for a given trait or set of traits that is due to selection, the use of which is increasingly being encouraged as a high-level performance indicator for plant breeding (Covarrubias-Pazarán, 2020). As a key indicator of biological (more specifically, quantitative genetic) change relative to selection practices, genetic gain bridges concerns over biological improvement of crops with concerns over the efficiency of breeding practice (thus reflecting a quest towards cost-efficiency comparable to that described by Curry, this volume, in relation to the rationalisation of genebanks). Previously to the introduction of genetic gain, breeding programmes have most commonly been evaluated by counting the number of varieties released, a measure which reflects neither the extent of trait improvement realised in new varieties nor their actual uptake among farmers. Genetic gain therefore provides an alternative and potentially more effective metric, increasingly used world-wide, to assess the success of breeding programmes and quantify the agronomic value of new varieties.¹

Alongside its use as an evaluative measure, normative commitments to increase the rate of genetic gain in breeding programmes have been established as key policy goals for plant breeding in recent years, for example in the funder-led Crops to End Hunger strategy of the CGIAR. The adoption of increased (or “accelerated”, as it is

¹The case study presented here draws on a more detailed analysis of genetic gain as an indicator, how it is being implemented in international plant breeding networks, and some key implications (see Williamson & Leonelli, 2022).

often phrased) genetic gain as a policy objective has been linked to: improved cost efficiency; better improvement of complex, quantitative crop traits; and the adaptation of agriculture to climate change through faster development of new varieties targeted to rapidly changing environments (Atlin et al., 2017). This objective has come with new reporting requirements for breeders and managers, and has been incorporated into formal systems for evaluating breeding programs such as the Breeding Program Assessment Tool, developed at the University of Queensland, evaluation through which is now mandatory for any programs receiving funding from the Bill and Melinda Gates Foundation.

Reporting on genetic gain is encouraged not only through the retrospective calculation of rates on the basis of historical data or designated ‘era’ trials using stored germplasm, but also through the estimation of future genetic gains from a given breeding programme design. Such estimations have significant implications for how breeding programmes are designed, with choices about breeding source materials, selection methods and trial environments, among other factors, decided on the basis of their estimated contribution to genetic gain.

There are several means through which rates of genetic gain can be increased. These include reducing the length of time that breeding takes (i.e. fewer breeding generations), improving selection accuracy, and/or increasing selection intensity (Williamson & Leonelli, 2022). The methods available to achieve these goals frequently involve complex forms of data linkage that have emerged since the turn of the millennium. Indeed, new methods to collate and integrate disparate data sources have arguably driven the turn towards viewing rapid improvement of genetic gain as a feasible goal, progress against which can be precisely measured and quantified. The most prominent example of this is Genomic Selection, whereby molecular marker data taken from biological samples can be used to predict the performance of individual plants based on their genotype, using complex and highly tailored models, thus allowing selection decisions to be made well before the plants in a given generation reach maturity (Xu et al., 2019). Other methods include the use of environmental characterisation together with climatic data (including predicted data) and crop modelling to increase the accuracy of selection by better targeting evaluation and selection to the environmental conditions for which a new variety is being bred (Ramirez-Villegas et al., 2020; Chenu, 2015). Across these methods, there is a particular emphasis on those that increase the speed of breeding, which is often a particularly cost-effective way of increasing genetic gain (Cobb et al., 2019: 634).

Increasing genetic gain thus plays directly into one vision of data-intensive agriculture – what we might call precision breeding – in which maximal trait improvements can be realised in a rapid space of time through tightly integrated pipelines for data collection, integration and analysis, moving back and forth between the field, sequencing labs and computational facilities (cf. Cobb et al., 2019). Rather than discussing genomic or environmental data linkage, which have been discussed at length elsewhere in this volume, this chapter will primarily focus

on the role and status of socio-economic data and knowledge in this data-intensive vision.² Understanding how such data is collected and used is critical to assessing the social dimensions of responsibility in plant and agricultural science. While privacy and data protection form one pillar of responsible practice in this domain, as discussed in the introduction to this volume and the chapter by Zampati, we are specifically thinking here about the possibilities that the integration of such data into plant breeding programs afford for achieving goals of socio-economic development and improved human wellbeing, for a diverse and inclusive constituency of actors. Our aim is to demonstrate how the conceptual and normative commitments that have accompanied the increased focus on maximising genetic gain, especially in the CGIAR, have significant implications for the kind of engagement with agricultural stakeholders that can be imagined and implemented. This in turn has implications for the kinds of social benefit that breeding programs can deliver.

The starting point for this discussion is a tension that has dogged appeals to accelerate genetic gain: Namely, that despite the range of potential benefits that have been attributed to this goal (such as improved yields or increased resilience of agricultural systems to climate change), it does not necessarily lead to greater adoption of new crop varieties by farmers (cf. Ceccarelli, 2015). Indeed, low rates of adoption of improved varieties among farmers in the Global South is a longstanding concern of breeders and managers working in international agricultural research (Atlin et al., 2017). In order to combat this problem, alongside calls to accelerate genetic gain there has been a recognition by breeders and funders that breeding needs to become more ‘demand-led’, responding more closely to the needs and desires of farmers and other actors in food systems, as part of a wider ‘varietal replacement strategy’ (Atlin et al., 2017; Cobb et al., 2019).³

Implementing demand-led breeding requires processes for accessing and utilising socioeconomic data that can inform breeding targets and selection decisions. The primary method being promoted for this task is product profiling. In the following section, we analyse what this involves and some of the conceptual implications and limitations that follow from it. We then discuss work that has been undertaken in recent years to overcome some of these limitations and ensure that product profiling is gender-responsive. Following this discussion of gender-responsive breeding, we compare data-intensive breeding methods based on product profiling to participatory methods. Using the example of the Mother and Baby Trial Design, we suggest that many of the principles and goals of gender-responsive breeding can be achieved more consistently and dynamically through the latter. Participatory methods have tended to be excluded from breeding programs focused on maximising genetic gain, however, in line with longstanding disputes about whether highly centralised

²On ‘visions’ or ‘imaginaries’ of socio-technical systems (including data systems) and their implications for science, society and the future, see Jasanoff and Kim (2015) and Leonelli (2021).

³Other solutions involve the use of policy levers to encourage or oblige seed companies and farmers to distribute and adopt new varieties, respectively (Spielman & Smale, 2017). These solutions have quite significant political, legal and economic implications regarding the control of seed systems, but we leave that topic aside for the purposes of this chapter (see Williamson & Leonelli, 2022).

breeding grounded in formal selection theory or decentralised, participatory breeding produce greater impact. We conclude the case study by looking at how new data infrastructures are being developed to facilitate dense data collection from participatory methods and their integration into breeding programs alongside other data-intensive methods such as Genomic Selection. We argue that these infrastructures point to alternative visions for breeding and agricultural development, but the prospects for wider adoption of such socially responsive, integrated programs will depend on the extent to which normatively entrenched goals such as accelerating genetic gain govern the distribution of resources and labour in international agricultural research.

2 Product Profiling and Gender-Responsive Breeding

Genetic gain is an indicator that can be assessed and realised for any given trait or index (set) of traits. While the selection of target traits for improvement has long been led by breeders, recently both public and private sectors have shifted to demand-led modes of breeding, where the choice of desirable traits is made through the collection and analysis of socio-economic data. A key method in that respect is product profiling (e.g. Persley & Anthony, 2017).⁴ This method has been strongly promoted by proponents of genetic gain, alongside changes to biological breeding methods. This is not only because it facilitates a demand-led approach to breeding, but also because it allows a rapid, formalised delivery of socio-economic information that can be integrated into the tight timescales and optimised pipelines needed to increase genetic gain (e.g. Cobb et al., 2019; Atlin et al., 2017).

A product profile can broadly be defined as “a set of targeted attributes that a new plant variety is expected to meet in order to be successfully released onto a market segment” (Ragot et al., 2018, cited in Cobb et al., 2019: 628). In other words, a product profile describes a plant variety viewed as a desirable replacement for already established varieties within a particular market, thus establishing a key objective for breeders’ work over the coming years. Indeed, product profiles are framed as a concise, formalised set of targets that can guide the design of a breeding programme and selection decisions throughout (cf. Ragot et al., 2018). They are assembled at the start of a breeding project by breeders in collaboration with market and socioeconomic researchers. Supporting the creation of product profiles are a set of techniques of market segmentation that allow the target constituency for a breeding programme to be identified and studied. These involve distinguishing distinct groups within a market, “segments” defined by “a relatively homogeneous demand for a commodity (here crop varieties or animal breeds)” (Gender &

⁴For a detailed case study of how product profiling methods are being incorporated into plant breeding (specifically cassava breeding) at one CGIAR centre, the International Institute of Tropical Agriculture (IITA), see Agbona et al. (this volume).

Breeding Initiative 2017, cited in Orr et al., 2018: 6). A target segment for breeding is then identified, based on the desired social and/or economic intervention of the breeding programme, and taking into account factors such as agroecosystems, demographics and technological skill. The target segment is the group (usually agricultural producers) who will adopt the resulting variety, although the actual beneficiaries of a breeding programme may be different, for example consumers of a food variety or other actors in the value chain. Product profiles can then be assembled by surveying the needs and desires of both the target segment and other stakeholders (see Orr et al., 2018; Ragot et al., 2018).

Product profiles are meant to facilitate demand-led breeding, where the demand primarily envisioned is market demand, focused on breeding crops that facilitate new commercial opportunities and advantages for farmers and other producers. This is a distinct and non-trivial commitment. While improving the economic position of farmers is a valuable goal that has ramifications for wellbeing and the fight against poverty, it is only one among several important objectives when considered from the wider perspective of social development, including climate action, responsible consumption and reduced inequalities (to pick the most relevant three objectives among the seventeen UN Sustainable Development Goals). The focus on market-led demand underpinning the construction of product profiles reflects a longstanding bias in development discourse and practice towards economic growth and commercialisation (cf. Escobar, 1995), a bias that was largely true of the Green Revolution (e.g. Harwood, 2020) and continues to be true of its legacy projects (e.g. Holt-Giménez, 2008). More ambitious sustainability-focused goals that don't necessarily contribute to market outcomes, such as supporting agroecological systems, tend to receive less support (cf. Rosset & Altieri, 2017).

This situation is problematic in several respects, and we shall here briefly discuss only one of them, concerning the intersection between product profiling, breeder communities and gender equality. It has been well documented that crop improvement focused on commercial value tends to favour men substantially more than women, especially in rural and underdeveloped agricultural settings (cf. Sachs, 2019). Gender differences provide an especially useful lens for thinking about social responsibility in relation to plant breeding, socioeconomic data and indicators such as genetic gain, so it is worth here turning to this topic in some detail.

In order to overcome some of the conventional biases in breeding towards forms of crop improvement that favour men, significant work has been undertaken to improve the gender-responsiveness of breeding in the CGIAR and related networks and institutions (for a history of this work, see Van der Burg, 2019, 2021). In recent years, this has included a significant push to design gender-responsive methods and principles for product profiling, organised through the CGIAR Gender and Breeding Initiative (e.g. Ashby et al., 2018; CGIAR Gender & Breeding Initiative, 2018). This work has been extensively documented by Ashby and Polar (2019), who also summarise some of the key differences in crop trait preferences and socio-economic position between men and women. Such differences are in practice highly variable, and there is no universal set of women's preferences as opposed to men's: In many cases, gendered preferences converge. Nevertheless, there are recurring themes that

can be used to guide the design of gender-responsive agricultural research and plant breeding practice (see also Sachs, 2019). One such theme is the importance often placed on particular qualities of the crop rather than overall yield. Due to the distribution of labour in the household economy, women will frequently prefer qualities that reduce labour (such as cooking time, or ease of peeling roots and tubers), even at some cost to overall yield. Where men might primarily be concerned with the income that can be made from selling a harvest in larger commercial markets, women frequently have to consider trade-offs related to household work, the sustenance of their own community and the ultimate end use of a crop (such as household processing and consumption), whether by themselves or by other local women to whom they might sell in more informal markets. As Ashby and Polar observe, it is necessary to consider “the different ways in which resources, rights and responsibilities are shared among women and men engaged in small-farm production, processing and marketing” (2019: 28–9). This is especially so because increased commercialisation resulting from the introduction of new, “improved” varieties can in practice lead to a loss of control for women as cultivation of those more lucrative crops are taken over by men (2019: 23).

Participants in the Gender and Breeding Initiative have made major contributions towards developing methods for incorporating “gender screening” into product profiling in order to take account of gendered differences, such as specific weighting techniques and the differentiation between “niche” and “game-changing” traits. As Ashby and Polar’s comment on the distribution of resources, rights and responsibilities indicates, understanding these matters requires in-depth socio-economic research on the relevant groups for whom breeding is targeted. This is where questions of data return to the fore. “There is a practical challenge, therefore”, they note: “how to systematize relevant information about gender differences, especially men’s and women’s trait preferences, in a way that breeders can factor it into their trait prioritization and product profiles” (Ashby & Polar, 2019: 13). Unfortunately,

much of the published information is inadequate for this task: it consists of a description of a trait preferred by women or ranked higher by women than by men, for example “earliness,” without an explanation of the desired extent or level of the trait. This limits the usefulness of the information to breeders, who need to understand what producers consider the desired performance level of a trait. Trait preferences are also reported without analysis of the socioeconomic characteristics of respondents other than their gender and geographic location. Simple sex-disaggregation of preference data is not very useful for informing breeding objectives, because it is essential to understand what resource constraints or producers’ objectives are associated with a given preference and whether there is an underlying gender inequality at work. In addition, data on gender differences in trait preference studies is too often reported without evidence that the respondents are representative of a clearly identified population of end users. This makes it difficult to draw general conclusions about the significance of a gender-differentiated trait preference at a scale that a breeding program can rely on, as predictive of widespread end-user acceptance. (2019: 29)

What this points to is the significant issues that remain around access to and integration of appropriately detailed socio-economic data and information. Indeed, as noted in the report on a CGIAR workshop on product profiling, “For some

questions, good evidence may not exist, and until it can be obtained, best instincts and knowledge from the breeding team may need to be used as a starting point” (CGIAR Gender & Breeding Initiative, 2018: 18).

In part, this situation relates to difficulties surrounding data collection. Indeed, Almekinders et al. (2019) have argued that methods for researching farmer seed demand present an under-acknowledged bottleneck to attempts to redesign breeding pipelines and to increasing adoption of improved varieties (cf. McEwan et al., 2021).⁵ Partly, however, it also relates to the structure of product profiling, which still places with breeders the responsibility for making critical decisions that have wide-ranging implications.

In order to arrive at a final Product Profile, breeders evaluate, weight and prioritize the individual plant traits under consideration for inclusion in the product profile. Trait prioritization is highly selective, because the number of traits that can be included in any one profile is usually restricted to prevent the selection process from becoming unduly complex. The criteria breeders use for trait prioritization are often a mix of commercial, technical and business considerations, shaped by the goals of the breeding program. (Ashby & Polar, 2019: 13)

Socio-economic data and expertise, including gender data, are only incorporated at such key decision points, and often through very informal means. This is understandable where information is a limited resource. Stepping back, however, we might throw this situation into relief by comparison with some alternative modes of breeding available, specifically those that take a more systematic approach to the inclusion of socio-economic data and knowledge through participatory approaches.

3 Participatory Breeding for Dynamic Socio-Economic Data Flows

Participatory plant breeding methods, involving farmers directly in the selection process for new varieties, began to emerge in the late 1970s before taking root more substantially in the 1990s (Harwood, 2012: 142–3; Cleveland & Soleri, 2002; Westengen & Winge, 2020). Participatory methods provide a very different model of socio-economic responsiveness in comparison to conventional, centralised breeding.

Consider the Mother and Baby Trial Design method developed for participatory potato selection at the International Potato Center (CIP) in Peru (De Haan et al., 2019). This method utilises a combination of a centrally managed, experimental field trial (the ‘mother’) in which multiple varieties are grown and smaller trial plots (the ‘babies’) in farmers’ fields that reflect the latter’s own agronomic conditions. Participating farmers engage in evaluation of the different varieties at key stages,

⁵The history of social research in the CGIAR has long been marked by highly variable investment and integration with core plant breeding and research activities (Cernea & Kassam, 2006).

from flowering through to harvest, at both the managed and on-farm plots. These evaluations include standard yield assessment, but more importantly they include evaluation on the basis of selection criteria that are identified and ranked by farmers themselves at the time of evaluation. These criteria may include relatively conventional trait preferences such as resistances to blight, but also trait preferences that are more contextual and tangential to crop production, such as the adequacy of foliage for feeding livestock (2019: 26–27). A particularly important set of additional evaluations are those concerning the qualities of the crop, especially qualities relating to cuisine and organoleptic (i.e. sensory) traits such as appearance, taste and texture (2019: 45–6). Once participant farmers have chosen their preferred criteria, they rank plant varieties on that basis through simple voting methods involving placing seeds or other tokens in paper bags.

Gender-responsiveness is critical to the Mother and Baby Trial Design. This is achieved, first, through the focus on crop qualities, which as we saw is typically favored by women participants; second, by ensuring that women have the space to make their own contributions and decisions free from the influence of male farmers; and third, by designing participation such that the data collected from these trials can be disaggregated by gender. Ensuring space for women may require not only an equal balance of female and male participants, but also conducting discussions and voting with women separately from men (2019: 26). The ability to disaggregate data by gender can be achieved by providing men and women with different seeds/tokens for voting that can be counted separately (2019: 27–8).

Participatory breeding methods such as the Mother and Baby Trial Design have the advantage of providing socio-economic data collection and integration that is more *consistent* and more *dynamic*: Consistent, because they do not depend on highly variable and often heavily mediated flows of information; and dynamic, because the data collected from and opinions offered by farmers contribute directly to the shaping and reshaping of breeding and selection decisions throughout the whole process. As Almekinders et al. note, “The picture we create of the farmers’ preferences is a snapshot taken from our perspective as researchers and devoid of trade-offs and considerations farmers have in a real-life situation” (2019: 17). This ‘snapshot’ quality is accentuated where socio-economic data is incorporated at a single decision point in the product profiling process.

On top of these advantages, and perhaps most critically, it has also been argued that participatory breeding leads directly to greater varietal adoption by farmers. Ceccarelli and Grando (2007) have observed that in conventional breeding “the entire process is supply-driven; as a consequence, in many developing countries many varieties are produced and released but only a small fraction of these are adopted. With [participatory plant breeding], decision[s] on which variety to release depend on initial adoption by farmers; the process is demand-driven” (2007: 356). This is quite a different model of demand-driven breeding to the idea of market demand discussed above, one in which demand is community-led and treated as demonstrable adoption by farmers rather than a ‘snapshot’ of preferences, thus building adoption itself into the breeding process. Moreover, Ceccarelli notes elsewhere that “in a conventional system, 5 to 6 [years] typically pass after official

release before appreciable adoption commences [. . .], and during this time, farmers' priorities, agronomic conditions (e.g., availability of irrigation or fertilizer price), policy measures (e.g., introduction or removal of subsidies), and market demands may change, making the breeding objectives set at the beginning of the breeding program obsolete" (2015: 89). The dynamic engagement with farmer needs, priorities and growing conditions in participatory breeding directly responds to such issues, ensuring that varieties remain relevant to changing conditions.

Given these advantages, then, why are participatory breeding methods practically invisible in key discussions of genetic gain (e.g. Atlin et al., 2017; Cobb et al., 2019), despite the corresponding concern for varietal adoption? And why are questions of the social responsiveness of breeding limited to those information flows that can be condensed into a limited set of goals captured in a product profile? This situation is not new. As Harwood notes, participatory breeding has often been strongly resisted by breeders, with many considering it "an unnecessary alternative to conventional breeding (rather than an additional option)" (2012: 146). Indeed, one prominent proponent of accelerating genetic gain has asked of participatory plant breeding, "Why do we need it? We need it because we don't do good market research to really understand what farmers need, what millers need, what consumers prioritise" (Atlin, 2016). The attention to additional actors in food systems beyond farmers is important. But as we have indicated above, it is often market research flows that tend to be inconsistent by comparison to participatory methods. Moreover, it is debatable whether much of what is conventionally conducted under the rubric of market research addresses socio-economic concerns over gender relations and the distribution of resources, rights and responsibilities, which Ashby and Polar (2019) among many others have flagged as vital to addressing social and economic inequalities.

More broadly, we take this discussion of product profiling in relation to participatory breeding methods as exemplifying the critical role of the conceptual and normative dimensions of plant breeding for the design and implementation of data-intensive approaches. Specifically, our analysis highlights a tension between how data-intensive plant breeding is being imagined and the practical requirements of organising participatory breeding schemes. When implemented within breeding programs, the commitment to maximise genetic gain is typically accompanied by a commitment towards speed and efficiency in the collation of data and criteria underpinning the choice of product profiles (e.g. Cobb et al., 2019: 634; cf. Williamson & Leonelli, 2022): the CGIAR for instance is pushing for tightly integrated pipelines for data production, integration and analysis, such that selection decisions can be brought forward and the length of time from initiation of breeding to variety release reduced, potentially by up to 5 years depending on the crop species and methods used. Product profiling is attractive in relation to these commitments, because it provides a clear and limited set of target traits for improvement that breeders can use to make selection decisions under conditions of time pressure, in conjunction with molecular and evaluation data drawn from field trials. In comparison, participatory breeding programs fare much worse: they require significantly higher investments to set up, especially if large numbers of farmers and on-farm trials are involved; and collection and analysis of data from those on-farm trials and

from participatory evaluation sessions takes considerable time, especially when compared to the possibilities of Genomic Selection to predict plant performance before it has even reached maturity. This can lead to drag on rates of genetic gain, by adding additional time and labour requirements into pipelines, making participatory breeding look unappealing despite the above-mentioned advantages in terms of supporting social equality and agrobiodiversity.

This in turn underscores a continuing tension between the commitment to accelerating genetic gain and the need to increase varietal adoption, goals which are practically and conceptually separated in current visions of data-intensive plant breeding. If the aim of public plant breeding is ultimately to deliver social as well as economic impact, then any accounting for the efficiency of breeding should factor in a combination of genetic gain, varietal adoption and agrobiodiversity assessment more broadly (cf. Ceccarelli, 2015). Focusing solely or even primarily on genetic gain and its delivery to farmers as key indicators of success for plant breeding risks perpetuating a situation of supply-driven breeding and market-led seed systems, where biotechnological improvement becomes a primary value and an end in itself, while the social impacts of breeding are shaped to accommodate this goal. When it comes to data-intensive breeding, it is not outlandish to suggest that responsible practice should invert this situation, with the social impact of breeding driving the choice and implementation of biotechnological improvement. We argue that this may require rethinking the maximisation of genetic gain as a *situated* rather than a universal objective: One that can be deployed in certain circumstances but should always take into account the potential conflicts this can produce with other commitments, rather than being imposed as a key objective across breeding programs at large and then onto seed systems, through a treadmill of variety release that is pushed onto farmers.⁶

4 Conclusion: Essential Components of Responsible Breeding Strategies

The eminent historian of agriculture James C. Scott has provided a provocative reading of efforts to improve agriculture through biotechnology, as follows: “if the logic of actual farming is one of an inventive, practiced response to a highly variable environment, the logic of scientific agriculture is, by contrast, one of adapting the environment as much as possible to its centralising and standardising formulas” (1998: 301). This controversial reading may be viewed as applying well to the current fixation on accelerating genetic gain, where the infrastructures and evaluative procedures supporting data-intensive breeding are constructed around highly centralised and standardised methods of product profiling, which do not admit –

⁶See footnote 2, and cf. Williamson & Leonelli, 2022 for a more detailed discussion of seed system issues.

through their commitment to speed and market-led understandings of varietal demand – of participatory approaches which may be slower and yet yield better outcomes in terms of social equality and support for agrobiodiversity. However, we do not think that it is necessary or even fully warranted to juxtapose conventional, data-intensive breeding focused on increasing genetic gain with participatory breeding methods, as if these two approaches were incompatible and intrinsically opposed to each other. What we have suggested is that there is a tension among some of the commitments explicitly or implicitly endorsed by these two approaches, which needs to be highlighted and critically discussed in order to successfully reconcile their respective advantages. In Scott's terms, there may be ways to reconcile the logic of actual farming with that of scientific agriculture, as long as a balance is sought between standardisation and speed on the one hand, and participation and inclusive data-intensive methods on the other.

This point has been made most thoroughly by Fadda et al. (2020), drawing on the example of the Bioversity International 'Seeds for Needs' project. What such projects indicate is not a necessary conflict between competing methods, "even though the two approaches are different from a conceptual and underlying philosophical point of views" (2020: 2), but the potential for an innovative and deeper *integration* of participatory methods with genomic and other data-intensive methods. Steps in this direction are also being taken by the cassava breeding programme at IITA, for instance through the use of the Tricot (triadic comparison of technologies) participatory methodology, which reflects similar goals to the Mother and Baby Trial Design method (see Agbona et al., this volume). In closing this chapter, we shall identify and discuss what we regard as three essential components to such an integrated approach to plant breeding.

The first component is **the development and reliable maintenance of digital infrastructures** that support the sourcing and integration of data from farmers and on-farm trials. This needs to include semantic standards that incorporate farmer and other local terminologies, such as the Crop Ontology (Arnaud et al., 2020; Leonelli, 2022). It also needs to include platforms for crowdsourcing participatory trial data directly from farmers, such as the ClimMob platform being developed to support the Tricot methodology (van Etten et al., 2020), which allow much greater scaling of participation, and thus greater efficiency and reliability of results (an aspect that has been the source of criticism by proponents of conventional breeding; e.g. Atlin et al., 2001). The appointment of 'quality champions' or similar designated experts to support the effective use of digital infrastructures, as has been undertaken for the BREEDBASE breeding data management system, also assists in addressing some of the critical organisational and skills issues that can limit the adoption of such technically and socially complex systems (Agbona et al., this volume). This is particularly effective when sourcing at least some experts from local communities. Here we see glimpses of future data-intensive plant science and related digital infrastructures being put directly in the service of social inclusion and responsiveness (similar to the blockchain schemes discussed by Kochupillai and Köninger, this volume). As other chapters in this volume indicate (e.g. Fullilove and Alimari), the possibilities for this being achieved in practice will depend heavily on institutional

norms and structures, and on whether concrete support – through policy and resource allocation – can be thrown behind such efforts. In any case, the significance of investment in reliable, well-maintained, long-term data infrastructures as a fundamental requirement for the sustainable use of data-intensive tools for plant breeding cannot be underestimated.

The second component encompasses the ability for plant breeders, data and plant scientists, farmers, policy-makers and industry representatives in this domain to **explicitly confront and discuss diverse assumptions relating to conservation, biodiversity and development**. Practically, this requires implementing processes through which diverse stakeholders can come together and engage one another in ways that make a meaningful difference to how research and development are done, as in the collaborative and open-ended forms of organisation that characterise the CoEx project discussed by Louafi et al. (this volume). This typically includes consultation with social scientists and local representatives that can broker diverse concerns and help identify and debate the underpinning conceptual and normative commitments of plant breeding strategies (whether current or imagined) and how those can be reconciled to foster responsible research practice within specific communities and locations.

What do we mean by conceptual and normative commitments? These are the scientific, social, economic and other foundational concepts mobilised in agricultural research and development, which may not be explicitly recognised yet underpin ongoing practices, including how breeding strategies and related forms of data linkage are being developed and implemented. These foundational concepts are often tacit or taken for granted, but have a wide range of implications. While the large-scale mobilisation of data provides new opportunities, our analysis of social responsiveness in genetic gain-focused breeding has highlighted how data-intensive visions of agricultural research can also produce frictions when located in the wider landscape of agriculture (cf. Edwards et al., 2011). Looking beyond this specific example, additional issues include: the uneven landscapes of both scientific understanding and data flows themselves, which create discrepancies and inequalities in the extent to which data-intensive methods can be applied and can work productively for different groups (Kochupillai and Köninger, this volume; Zampati, this volume); the conceptual and cultural gulf between farming communities and research scientists when it comes to agricultural strategies (Louafi et al., this volume); and indeed the lack of training for scientists themselves to recognise and understand alternative narratives of agricultural development (and where data science can fit in these).

This is important for responsible research practice in plant data linkage for at least three reasons. First, because unquestioned, dogmatic adherence to specific normative commitments can lead to aspects of research practice becoming centralised and entrenched (materially as well as culturally) as the necessary or right way for things to be done, and block off alternatives (Scott, 1998). Second, because scientific research does not just exist in its own bubble; it feeds into much wider imaginaries of society, economy, development, and so on, which in turn also influence the ways we imagine and conceptualise science (Jasanoff, 2004). And third, because the extensive and highly diversified impact of plant breeding and agronomic strategies

on planetary health makes it imperative to continue to look for alternatives and/or localised solutions, both for how science is done and for agricultural development, and to consider whether such alternative and/or localised approaches may improve current practice.

Following this, the third component we identify as crucial to an integrated and responsible approach to plant breeding is **interdisciplinary and transdisciplinary collaboration**, particularly involving historical, philosophical and social studies of science, to consider critically the implications of entrenching concepts into infrastructures – and possible alternatives. Within this volume, many examples have been given of ways to broker social and scientific considerations within data-intensive breeding. Most chapters have pointed towards ways to remain evidence-based and build on innovative data-intensive tools, while at the same time grounding novel forms of data linkage on an understanding of the geographically and conceptually diverse histories of agricultural policies and technologies. Among the many examples of such work available beyond this volume, one of the most relevant is Jonathan Harwood's (2012) effort to uncover a forgotten history of public plant breeding in southern Germany, predating the Green Revolution. Harwood uses the case to think about issues of who is supported by agricultural research and development and in what ways, particularly through a comparison with Green Revolution breeding and the growth and decline of participatory breeding methods in the CGIAR throughout the 1990s (an example that resonates with the case we have presented above).⁷ An additional example is the recent Nuffield Council on Bioethics (2021) report on genome editing and farmed animal breeding, which draws on expertise from a range of disciplines across the biosciences, social sciences and humanities. Reflecting the concerns in this chapter for how data are assembled and indicators put to work in breeding practice, the authors analyse the scope and purposes of indices used to evaluate breeding animals. Among the recommendations made in the report are the need to expand the scope of the indices to include traits of public or social as well as economic value, for example those related to health traits or traits that can impact climate emissions (2021: 155–160, 192–3). The kinds of conceptual and normative considerations raised in these examples, and throughout this chapter, can crucially inform research and policy decisions around how to set up infrastructures, data governance and institutional goals for agricultural development and food security. Responses are likely to involve elements of design of socio-technical systems, thus intersecting strongly with the design of technical infrastructures whose significance we just emphasised.

In closing, it is important to stress that consideration of responsibility and social responsiveness introduced through a focus on the conceptual and normative dimensions of plant and agricultural data linkage does not produce clear, unambiguous conclusions. Insights tend to be context-specific, and thus require detailed attention to and knowledge about how research and development is set up in practice.

⁷Harwood has also made complementary arguments about the relevance of history of science and development to policy (Harwood & Sturdy, 2010; Harwood, 2018).

Historical, philosophical and sociological studies of science provide excellent background knowledge on these aspects; but they need to be complemented by practical and tacit knowledge held by domain experts – an interdisciplinary dialogue that this volume has attempted to contribute towards establishing. Moreover, tensions and disagreements are unlikely to be resolved easily, with disagreements over the relative value of centralised, formal breeding methods versus decentralised, participatory methods running for several decades now. In data-intensive science as in other realms of research, responsibility involves opening up such matters to public debate and the option of co-producing future strategies with relevant stakeholders and publics.

References

- Almekinders, C. J. M., Beumer, K., Hauser, M., Misiko, M., Gatto, M., Nkurumwa, A. O., & Erenstein, O. (2019). Understanding the relations between farmers' seed demand and research methods: The challenge to do better. *Outlook on Agriculture*, 48(1), 16–21. <https://doi.org/10.1177/0030727019827028>
- Arnaud, E., Laporte, M. A., Kim, S., Aubert, C., Leonelli, S., Cooper, L., Jaiswal, P., Kruseman, G., Shrestha, R., Buttigieg, P. L., Mungall, C., Pietragalla, J., Agbona, A., Muliro, J., Detras, J., Hualla, V., Rathore, A., Das, R., Dieng, I., et al. (2020). The ontologies community of practice: An initiative by the CGIAR platform for big data in agriculture. *Patterns*, 1, 100105. <https://doi.org/10.1016/j.patter.2020.100105>
- Ashby, J. A., & Polar, V. (2019). The implication of gender relations for modern approaches to crop improvement and plant breeding. In C. E. Sachs (Ed.), *Gender, agriculture and agrarian transformations: Changing relations in Africa, Latin America and Asia*. Routledge.
- Ashby, J. A., Polar, V., & Thiele, G. (2018). *Critical decisions for ensuring plant or animal breeding is gender-responsive* (CGIAR Gender & Breeding Initiative, Brief 1). CGIAR.
- Atlin, G. (2016, April 29). *Modernizing plant breeding programs to deliver higher rates of genetic gain in the developing world*. ICRISAT seminar. Available at: <https://youtu.be/r2Zh64QD8TQ>
- Atlin, G. N., Cooper, M., & Bjornstad, A. (2001). A comparison of formal and participatory breeding approaches using selection theory. *Euphytica*, 122, 463–475. <https://doi.org/10.1023/A:1017557307800>
- Atlin, G. N., Cairns, J. E., & Das, B. (2017). Rapid breeding and varietal replacement are critical to adaptation of cropping systems in the developing world to climate change. *Global Food Security*, 12, 31–37. <https://doi.org/10.1016/j.gfs.2017.01.008>
- Ceccarelli, S. (2015). Efficiency of plant breeding. *Crop Science*, 55, 87–97. <https://doi.org/10.2135/cropsci2014.02.0158>
- Ceccarelli, S., & Grando, S. (2007). Decentralized-participatory plant breeding: An example of demand driven research. *Euphytica*, 155(3), 349–360. <https://doi.org/10.1007/s10681-006-9336-8>
- Cernea, M. M., & Kassam, A. H. (Eds.). (2006). *Researching the culture in agri-culture: Social research for international development*. CABI Publishing.
- CGIAR Gender and Breeding Initiative. (2018, November 12–13). *Gender-responsive product profile development tool*. Workshop report. Available at: <https://cgspace.cgiar.org/handle/10568/99094>
- Chenu, K. (2015). Characterizing the crop environment – Nature, significance and applications. In V. O. Sadras & D. F. Calerini (Eds.), *Crop physiology: Applications for genetic improvement and agronomy* (2nd ed.). Academic Press.

- Cleveland, D. A., & Soleri, D. (Eds.). (2002). *Farmers, scientists and plant breeding: Integrating knowledge and practice*. CABI Publishing.
- Cobb, J. N., Juma, R. E., Biswas, P. S., Arbelaez, J. D., Rutkoski, J., Atlin, G., Hagen, T., Quinn, M., & Hwa Ng, E. (2019). Enhancing the rate of genetic gain in public-sector plant breeding programs: Lessons from the breeder's equation. *Theoretical and Applied Genetics*, 132, 627–645. <https://doi.org/10.1007/s00122-019-03317-0>
- Covarrubias-Pazarán, G. E. (2020). *Genetic gain as a high-level key performance indicator*. CGIAR excellence in breeding platform, breeding scheme optimization manual. <https://excellenceinbreeding.org/toolbox/tools/eib-breeding-scheme-optimization-manuals>
- De Haan, S. E., Salas, C., Fonseca, M., Gastelo, N., Amaya, C., Bastos, V. H., & Bonierbale, M. (2019). *Participatory varietal selection of potato using the mother & baby trial design: A gender-responsive trainer's guide*. International Potato Center. <https://cgspace.cgiar.org/handle/10568/106633>
- Edwards, P. N., Mayernik, M. S., Batcheller, A. L., Bowker, G. C., & Borgman, C. L. (2011). Science friction: Data metadata and collaboration. *Social Studies of Science*, 41(5), 667–690. <https://doi.org/10.1177/0306312711413314>
- Escobar, A. (1995). *Encountering development: The making and unmaking of the Third World*. Princeton University Press.
- Fadda, C., Mengistu, D. K., Kidane, Y. G., Dell'Acqua, M., Pe, M. E., & van Etten, J. (2020). Integrating conventional and participatory crop improvement for smallholder agriculture using the seeds for needs approach: A review. *Frontiers in Plant Science*, 11, 559515. <https://doi.org/10.3389/fpls.2020.559515>
- Harwood, J. (2012). *Europe's green revolution and others since: The rise and fall of peasant-friendly plant breeding*. Routledge.
- Harwood, J. (2018). Another Green Revolution? On the perils of 'Extracting lessons' from history. *Development*, 61, 43–53. <https://doi.org/10.1057/s41301-018-0174-5>
- Harwood, J. (2020). Whatever happened to the Mexican Green Revolution? *Agroecology and Sustainable Food Systems*, 44(9), 1243–1252. <https://doi.org/10.1080/21683565.2020.1752350>
- Harwood, J., & Sturdy, S. (2010). What can development policy learn from the history of development? *Food Security*, 2, 285–290. <https://doi.org/10.1007/s12571-010-0067-2>
- Holt-Giménez, E. (2008). Out of AGRA: The Green Revolution returns to Africa. *Development*, 51(4), 464–471. <https://doi.org/10.1057/dev.2008.49>
- Jasanoff, S. (Ed.). (2004). *States of knowledge: The co-production of science and social order*. Routledge.
- Jasanoff, S., & Kim, S.-H. (Eds.). (2015). *Dreamscapes of modernity: Sociotechnical imaginaries and the fabrication of power*. University of Chicago Press.
- Leonelli, S. (2021). Data science in times of pan(dem)ic. *Harvard Data Science Review*, 3. <https://doi.org/10.1162/99608f92.fbb1bdd6>
- Leonelli, S. (2022). Process-Sensitive Naming: Trait Descriptors and the Shifting Semantics of Plant (Data) Science. *Philosophy, Theory and Practice in Biology*. <https://doi.org/10.3998/ptpbio.16039257.000000>
- McEwan, M. A., Almekinders, C. J. M., Andrada-Piedra, J. J. L., Delaquis, E., Garrett, K. A., Kumar, L., Mayanja, S., Omondi, B. A., Rajendran, S., & Thiele, G. (2021). "Breaking through the 40% adoption ceiling: Mind the seed system gaps." A perspective on seed systems research for development in One CGIAR. *Outlook on Agriculture*, 50(1), 5–12. <https://doi.org/10.1177/0030727021989346>
- Nuffield Council on Bioethics. (2021). *Genome editing and farmed animal breeding: Social and ethical issues*. Report. Nuffield Council on Bioethics <https://www.nuffieldbioethics.org/publications/genome-editing-and-farmed-animals>
- Orr, A., Cox, C. M., Ru, Y., & Ashby, J. (2018). *Gender and social targeting in plant breeding*. CGIAR Gender & Breeding Initiative, working paper 1. <http://www.rtb.cgiar.org/gender-breeding-initiative/resources/>

- Persley, G. J., & Anthony, V. M. (Eds.). (2017). *The business of plant breeding: Market-led approaches to new variety design in Africa*. CABI.
- Ragot, M., Bonierbale, M., & Weltzein, E. (2018). *From market demand to breeding decisions: A framework*. CGIAR Gender & Breeding Initiative, working paper 2. <http://www.rtb.cgiar.org/gender-breeding-initiative/resources/>
- Ramirez-Villegas, J., Milan, A. M., Alexandrov, N., Asseng, S., Challinor, A. J., Crossa, J., van Eeuwijk, F., Ghanem, M. E., Grenier, C., Heinemann, A. B., Wang, J., Juliana, P., Kehel, Z., Kholova, J., Koo, J., Pequeno, D., Quiroz, R., Rebolledo, M. C., Sukumaran, S., et al. (2020). CGIAR modelling approaches for resource-constrained scenarios: I. Accelerating crop breeding for a changing climate. *Crop Science*, 60(2), 547–567. <https://doi.org/10.1002/csc2.20048>
- Rosset, P. M., & Altieri, M. A. (2017). *Agroecology: Science and politics*. Practical Action Publishing.
- Sachs, C. E. (Ed.). (2019). *Gender, agriculture and agrarian transformations: Changing relations in Africa, Latin America and Asia*. Routledge
- Scott, J. C. (1998). *Seeing like a state: How certain schemes to improve the human condition have failed*. Yale University Press
- Spielman, D. J., & Smale, M. (2017). *Policy options to accelerate variety change among small-holder farmers in South Asia and Africa South of the Sahara*. IFPRI discussion paper 01666. <https://www.ifpri.org/publication/policy-options-accelerate-variety-change-among-small-holder-farmers-south-asia-and-africa/>
- Van der Burg, M. (2019). “Change in the making”: 1970s and 1980s building stones to gender integration in CGIAR agricultural research. In C. E. Sachs (Ed.), *Gender, agriculture and agrarian transformations: Changing relations in Africa, Latin America and Asia*. Routledge
- Van der Burg, M. (2021). Gender integration in agricultural research for development. In C. E. Sachs, L. Jensen, P. Castellanos, & K. Sexsmith (Eds.), *Routledge handbook of gender and agriculture*. Routledge
- van Etten, J., Abidin, E., Arnaud, D., Brown, E., Carey, E., Laporte, M.-L., López-Noriega, I., Madriz, B., Manners, R., Ortiz-Crespo, B., Quirós, C., de Sousa, K., Teeken, B., Tufan, H.A., Ulzen, J., & Valle-Soto, J. (2020). *The tricot citizen science approach applied to on-farm variety evaluation: Methodological progress and perspectives*. CGIAR Research Program on Roots, Tubers and Bananas (RTB). RTB Working Paper. No. 2021-2.
- Westengen, O. T., & Winge, T. (Eds.). (2020). *Farmers and plant breeding: Current approaches and perspectives*. Routledge
- Williamson, H. F., & Leonelli, S. (2022). Accelerating agriculture: Data-intensive plant breeding and the use of genetic gain as an indicator for agricultural research and development. *Studies in History and Philosophy of Science*, 95, 167–176. <https://doi.org/10.1016/j.shpsa.2022.08.006>
- Xu, Y., Liu, Z., Fu, J., Wang, H., Wang, J., Huang, C., Prasanna, B. M., Olsen, M. S., Wang, G., & Zhang, A. (2019). Enhancing genetic gain through genomic selection: From livestock to plants. *Plant Communications*, 1(1), 100005. <https://doi.org/10.1016/j.xplc.2019.100005>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

