WILL ORGANIZATION DESIGN BE AFFECTED BY BIG DATA?

GILES SLINGER • RUPERT MORRISON

Abstract: Computing power and analytical methods allow us to create, collate, and analyze more data than ever before. When datasets are unusually large in volume, velocity, and variety, they are referred to as "big data." Some observers have suggested that in order to cope with big data (a) organizational structures will need to change and (b) the processes used to design organizations will be different. In this article, we differentiate big data from relatively slow-moving, linked people data. We argue that big data will change organizational structures as organizations pursue the opportunities presented by big data. The processes by which organizations are designed, however, will be relatively unaffected by big data. Instead, organization design processes will be more affected by the complex links found in people data.

Keywords: Organization design, big data, organizational structure, organization design process

We participated in the Big Data and Organization Design conference in Paris, May 2013, as representatives of Concentra, our consulting firm which specializes in design, data analytics, and technology. Many speakers at the conference discussed the various impacts of big data, defined as "high-volume, -variety, and -velocity information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making" (Laney, 2001). With respect to business organizations, big data allows more accurate customer segmentation in marketing. In health care, big data supports more targeted diagnosis and treatment. In employee recruiting, big data allows employers to screen more accurately. In the supply chain, big data reduces inventory wastage. And big data promises to alter the shape of organizations in many as yet unknown ways.

At the enterprise level, Galbraith (2012) proposes that big data will change organization structure, as large multinational firms will restructure to add a (fifth) structural dimension. Some new functions will engage in big data operations, which will distinguish themselves from the rest of the organization, just as the previous four structural emphases divided the organization according to (1) functions, (2) product divisions, (3) international units, and (4) customer segments. Big data will also change roles and power structures (Galbraith, 2014).

At the moment, big data is a promising technological innovation that may affect many business models. Is it really a step change, however, in its effects on how organizations are structured? Further, is big data going to change the processes by which we design organizations? Based on our experience from working with clients on organization design projects, we believe that big data will affect how organizations are structured more than how they are designed. In this article, we explore three propositions:

1. Organizations will be restructured to take advantage of big data opportunities.
2. Processes of organization design are unlikely to change because of big data.
3. The organization design process is not based on the volume, variety, and velocity of data; it is based on the slow-moving, linked nature of people data.

Journal of Organization Design
JOD, 3(3): 17-26 (2014)
DOI: 10.7146/jod.9729

ORGANIZATIONS WILL BE RESTRUCTURED BECAUSE OF BIG DATA

This proposition looks like an open-and-shut case – organizations are already restructuring to deal with big data. Galbraith (2014) discussed the example of Proctor & Gamble, which has created "control towers" to maintain continuously updated control of its supply chain. He also highlighted Amazon, which says it wants to be the world's most customer-centric organization, mostly by understanding its customers' data in great depth. And he described Nike, which created its Nike Digital Sports division in 2010, putting sensors in shoes, clothes, and watchbands, and setting up virtual athletics communities.

Why should companies restructure themselves to deal with big data? Resource allocation becomes much more flexible in organizations that can apply big data. With visibility of demand levels and supply volumes, they find it easier to move people, capital, and other resources across sites, functions, roles, and positions. For example, a theme park can reallocate staff quickly to busy areas, or a supermarket can respond rapidly to forecasts of changing weather conditions. Today, we find it normal that a supermarket chain should seek detailed insights on the impact that weather conditions have in different store locations and at different times on customers' behavior. Yet analyses as recent as Starr-McCluer (2000) could find only modest impact of weather on sales. Such an analysis seems almost to be from a different era because data available at that time was only available in aggregated forms. Starr-McCluer's data sources, for example, were average monthly temperature data across the whole of the U.S. and average monthly sales data across ten types of retail operations. This compares with modern sales data, which are minute-by-minute, and modern weather metrics, which are hour-by-hour, both types of data allowing for real-time analytics and decision-making.

It is clear that the volume and granularity of big data opens up possibilities that have never previously existed to track the supply chain and customer-company interaction. That will mean opportunities to deliver better services which, in turn, will require different kinds of organizational structures. Interestingly, Galbraith (2012) argues that this will also generate new tensions, and the most successful organizations will be those that manage the conflicts of direction and interest that will inevitably arise from having up to five different structural emphases in the business. Galbraith's (2012) observation echoes that of McAfee and Brynjolfsson (2012), who point out that big data will alter the sources of influence in the organization. The location of decisions will change, for example, as HiPPOs (the "Highest Paid People in the Organization") find that they need to allow their judgment-based decision-making to be modified, and at times overruled, by data-driven insights.

PROCESSES OF ORGANIZATION DESIGN ARE UNLIKELY TO CHANGE BECAUSE OF BIG DATA

Proposition 2 is much harder to affirm than Proposition 1. The overall process of organization design is typically described in stages, moving from the outside in: (a) environmental (external) analysis, (b) definition of the organization's purpose and mission, (c) assessment of the existing organization (internal analysis), (d) detailed design, and (e) implementation and review. An example of the overall design process is shown in Figure 1.

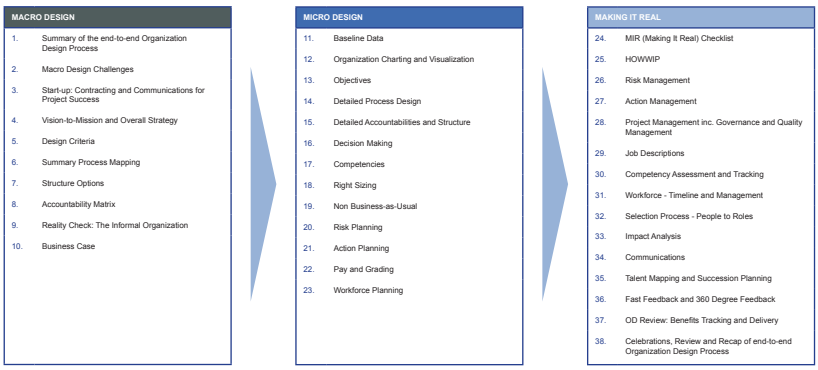

Fig. 1. Organization Design Process

Will the processes of organization design, as illustrated in Figure 1, change as a result of big data? Our firm has asked many academics and businesspeople how they think big data will change the process of designing organizations. A few predictions have emerged – crowd-sourcing of ideas for process change (a new source of data), data-mining CV or LinkedIn text for information on competencies (a new use of unstructured data), fast adjustment of organizational objectives in response to changing market conditions (a new use of high-velocity data) – but no clear patterns are evident. Therefore, we gathered examples of big data's impact that we could find in the literature and assessed whether each one would change the *structure* of organizations or the *process* of designing organizations (see Table 1).

Table 1. How Big Data Affects Organizational Structure and Design Process

What Makes it Big data?	Example	Change Organizational Structure?	Change Design Process?	Comments
Volume of data	Crowdsourcing of ideas for change in products or processes	Potentially by altering the products or services provided	No	Data mostly used for product and service re-design. Could indirectly impact the shape of organization needed, e.g., Sainsbury's: (a) Colleague Feedback panel has 3000 members; (b) "Tell Justin" gave >30,000 ideas during 2006–2010.[1]
Variety of data	Data on behavior, capabilities, personality profile, performance, absence, ENPS, NPS, mood, text, image data, etc.	Yes—could affect allocation of people to roles	Does not affect the method for designing roles	Many data types used historically for individual performance assessment and development are now available for analytics of large groups.
Velocity of data (1)	Staffing in response to changing external demand levels	Yes—alters the number of roles in the organization continuously	No—the structures are designed in the same way	The organization does not have to restructure itself formally. It uses a more effective right-sizing process to allocate staff where needed.
Velocity of data (2)	Internal fast feedback, as opposed to annual surveys	No—is about quality control, not structural change	Faster data—if linked to structures, clients or skills—can make responses more rapid but not different in type	Monthly data on management performance allows more rapid intervention. Fast feedback can help the organization respond to managers' training needs, but does not affect the organization structure.
Velocity of data (3)	Ability to respond in real time to customer needs or security issues	Only if organization needs new structures to respond	No—the design process is unchanged	Structural change required only if organization cannot "increase its clock speed" through current structural forms.

Our conclusion, based on assessing examples from the literature, is that the conventional big data factors (volume, variety, velocity) will affect how organizations are structured but not the process by which they are designed. To support this assertion, we documented the sizes of the largest datasets that Concentra has used in its various organization design projects (see Table 2). It is notable that the "macro" design stage – which is often the focus of organization design theory – generally uses smaller datasets. It is only for the very largest organizations, and for very detailed feedback or planning, that dataset sizes are larger than a million data points. Larger datasets occur in the "micro" design and implementation stages when the organization is making detailed evaluations of accountabilities, objectives, decisions, and competencies associated with individual employees.

1 Allen (2010) and Transparent Consulting interviews with Sainsbury's HR team, 2010.

Table 2. Data Requirements for 20 Organization Design Activities, Sorted by Stage of Implementation

Design Activity	Phase	Data Used	Data Points Per Size of Org (FTE)		
			1,000	10,000	100,000
Start-up: Contracting and Communications for Project Success	Macro Design	For informal network mapping: n employees * at most (5 * influence ratings from syrveys + 12 * email traffic mapping outputs + 3 * mapped expert references)	20,000	200,000	2,000,000
High-Level Process Design	Macro Design	List of outputs and 100-10,000 processes	100	1,000	10,000
Vision-to-Mission and Overall Strategy	Macro Design	List of company's stated mission, strategy items	20	20	20
Design Criteria	Macro Design	5-10 design criteria	10	10	10
Structure Options	Macro Design	3-6 structure options	10	10	10
Objective Mapping	Micro Design	n employees * 10 objectives * 12 datapoints	120,000	1,200,000	12,000,000
Detailed Accountabilities and Structure	Micro Design	n employees * 10 items responsible + 50 items supporting + 10 items approving	70,000	700,000	7,000,000
Decision Making	Micro Design	n employees * up to 50 decisions	50,000	500,000	5,000,000
Organization Charting and Vissalization	Micro Design	ID, Manager ID * n employees	2,000	20,000	200,000
Baseline Data	Micro Design	Summary business data - headount, orles, key outputs, objectives, KPIs	22	22	22
Right Sizing via Benchmark Comparisons	Micro Design	Up to 100 key comparisons vs. own data (e.g. % headcount in core function)	200	200	200
Fast Feedback	Implementation	Monthly feedback on 5 questions on implementation effectiveness for up to 3 years from n employees	180,000	1,800,000	18,000,000
Selection Process - People to Roles	Implementation	Mapping of 50 skills vs. 5 roles for n/10 candidates	25,000	250,000	2,500,000
Competency Assessment and Tracking	Implementation	n employees * 10-50 competencies * periodic updates	20,000	200,000	2,000,000
Impact Analysis	Implementation	Mapping of impact of To-Be vs As-I on e.g. 5 dimensions for each employee (e.g. line manager, location, role, pay, hours)	10,000	100,000	1,000,000
Talent Mapping	Implementation	Map n employees into 9 categories (for 9 box grid)	9,000	90,000	900,000
Succession Planning	Implementation	Map up to 5 employees into each role needing successor	5,000	50,000	500,000
Workforce Planning - Timeline and Management	Implementation	Plan for up to 1000 roles* periodic headcount per role (e.g. monthly for 3 years), plus tracking actual vs. plan	720	7,200	72,000
Pay and Grading	Implementation	Analysis of job complexity on 10 dimensions * number of distinct roles	100	1,000	10,000
Job Descriptions	Implementation	10-1000 standard role descriptions	10	100	1,000

Source: Concentra Consulting, OrgVue, Slinger (2014)

In short, it appears that the scale of data is not the major challenge in the processes used for designing organizations. Nor is it the speed of change in the data. Instead, it seems that the primary data challenge in the design process is how to deal with the slow-moving, linked nature of people data.

THE PROCESS OF ORGANIZATION DESIGN WILL BE BASED ON SLOW-MOVING, LINKED PEOPLE DATA

To understand the organization design process, we believe it is useful to focus on a particular type of data: people data. We define people data as data that has the worker – the current, potential, or former employee or contractor – as a key unit of analysis. People datasets are more often available, and they are richer today than in the past. Employers can, in principle, collect extensive information on daily productivity, working time, location, and even e-mail exchanges and other forms of social interaction. However, the most common elements of people data used for organization design are the same as they have been for a long time: current and forecasted headcount, fully loaded personnel costs, skills and experience, project preferences, and so on.

Why the Individual Matters as a Unit of Analysis

Using the individual as the unit of analysis puts a constraint on a full optimizing approach to organization design. In principle, organizations should be designed by "pull" – as a flow from the products and services that customers will buy, through the activities required, through the competencies needed, through the roles that cluster the competencies, to the teams that bring together the roles, and lastly, to the organizational structures that bring the teams together. The model that could exist in principle, however, hits a conceptual stumbling block. Competencies do not link directly to roles. They are clustered in people. People do not change very fast, and they are not divisible. Working hours and salaries are, by law, difficult to alter. New skills take time to learn. Relationships between peers and across hierarchical levels take time to build. Creativity and commitment influence the quality of output. In sum, the embodiment of organizational characteristics in units of people changes the nature of the design problem.

People Data Are Not "Big"

People data have always posed challenges for organizational analysis, but we should not overestimate that analytical challenge. People data are sparsely populated and slow moving. The actual number of links between organizational components is low. For example, imagine an organization of 1,000 people in 1.1 roles each, ten activities each, and ten products and twenty clients handled by each. The organization has, at most, a total of 2.2 million connections – still a lot in absolute terms, but not an especially big number for analytical purposes. The data in this "small" people dataset are usually incomplete, changing, and linked in complex ways that makes the organization design process challenging.

Dealing with People Data: Iteration and Simplification

Organization designers have responded to the challenge of optimizing the performance of the organization as a complex system by iteration and simplification. The iterative approach reminds us to treat the organization as a system. Interdependencies and linkages within the organization mean that change must be tested and cascaded layer by layer. This approach has the benefit of reducing risks caused by unexpected complex interactions within the system but will result in local variations from the preferred overall design. Some of those variations may be appropriate, and some may be costly, but no systemically optimal design exists. The simplifying approach may view the overall organization as a system, but typically it intervenes on one aspect such as demographics, competencies, talent, succession, or activity costs. This can provide consistency of treatment across the organization (e.g., standard processes, standard ratios, standard pay rates, standard spans of control), but the resulting design may be susceptible to unexpected consequences. Both approaches are adaptations to deal with the challenge of optimal organization design, but neither approach models the organization as a system.

How People Data Affect the Design Process

People data affect both the organization's design and the process of design itself. By employing some of the newer uses of people data, designers can come closer to optimizing their designs. We discuss some of these newer uses below, and they are summarized in Table 3.

Table 3. Where People Data Might Affect the Organization Design Process

Uses of People Data	Example	Impact on Organization Structure	Impact on Design Process	Implications
Informal Networks Data	Data on social capital—influence, communications, importance to the organization through informal networks	Yes—could lead to simpler structures if it showed formal hierarchies not needed	Yes—supporting an evolving organization design with leaders emerging through interaction	Extends mapping of the organization into data on new types of relationships—not just hierarchical or matrixed, but informal and influence. Offers additional "capability" variable(s) per person.
Data Visualization	People data particularly relevant for expressing via color, size, shapes, and hierarchical structures	No direct impact on structure	Adjustment more likely if managers see where costs, skills, and customer impacts are	Visualization could affect organization design—by giving a sense of the organization more intuitively, it might be more possible to achieve an organization design that makes sense to more people.
Incomplete Data	People data might be incomplete, but might still be necessary for organization design	Could allow simpler and more flexible structures	Incomplete data could be used for incomplete design—x% adaptable	Organizations have always been re-designed on the back of napkins. However, it would be innovative if organizations were designed consciously to cope with incomplete data.
Data Reflexivity	People datasets can affect themselves—as expressed in the feedback loops in Silverman Research's Social Media Garden	Does not change structure directly	Yes—the organization design process can evaluate its own progress and adapt	An exciting extension of the idea of group training environments where the group is explicitly invited to reflect on its own process, take ownership of it, and improve how it operates (Silverman, 2012).
Linked Relationships	People data are unusually highly linked—to processes, costs, customers, skills, services, objectives, etc.	Potential new role for strategy team / MI team	Yes—design and monitor systemic impact during the change process	It has always been hard to link and process the data. As this becomes possible on an ongoing basis, people will be more able to reconfigure their organizations as needed.

Informal networks data. Stephenson and Lewin (1996), Farmer (2008), Cross (2009), and others have investigated the informal networks that exist within and outside formal hierarchical structures. For example, an individual's influence (or social capital) can be mapped onto networks of innovation, knowledge, and collaboration, amongst others. Informal network analysis can be used during an organizational redesign as a reality check by asking, for instance, are our nominated leaders really people of influence? An example of informal networks analysis is shown in Figure 2, where managerially nominated influencers who may hitherto have been "over-recognized" (red) are contrasted with peer-nominated influencers who may have been "under-recognized" (green).

Informal networks data can also be used to plan for how an organization will evolve flexibly, as in the design of the U.K. government's transport innovation network (Transport Catapult, 2014). It has been designed to include innovation teams with individuals selected both for their capabilities and for their personal innovation networks. Collaboration levels are measured monthly on internal team working, inter-team working, and interactions with key innovators in parent organizations. Project manager roles are filled by emergent leaders from within each project team, typically within the first six weeks of the project team's life. The "health" of these networks is a key leading indicator of innovation success or failure (Farmer, 2013).

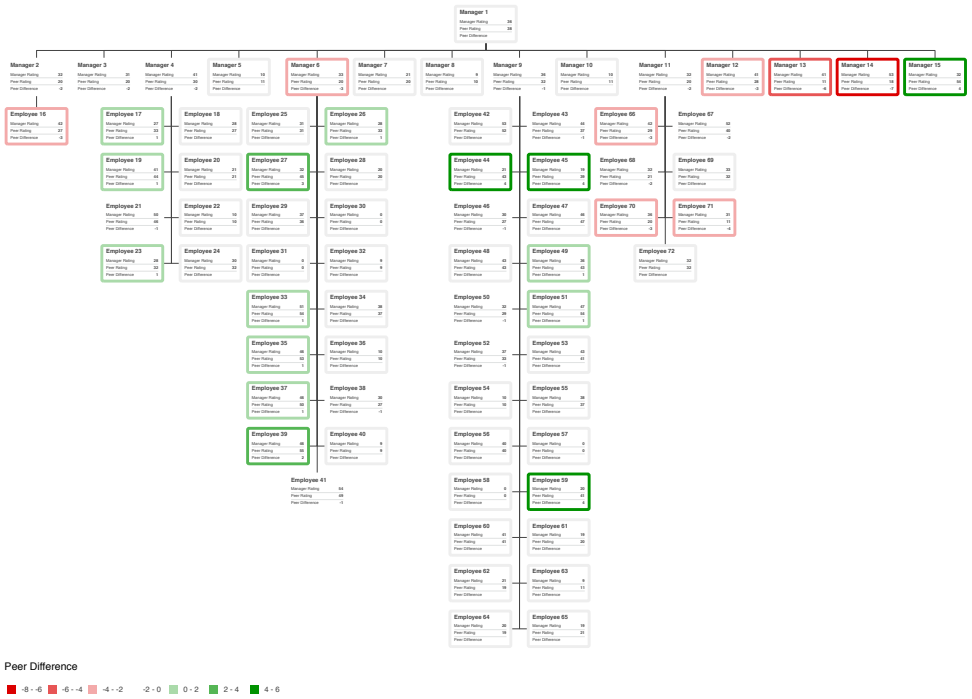

Fig. 2. Mapping Informal Networks on Hierarchical Structures

Data visualization. Davenport and Patil (2012) have argued that data cleansing, organization, and visualization will be critical skills for managing big data. Our recent consulting work has shown us that visualization also works well for people data. Figure 3 shows a business firm which had historically visualized its cost in tables of numbers or bar charts per division. It mapped people to processes using linking software to see the cost per process for everything it did. The impact was that at a single click the organization could change between seeing itself as a hierarchy and seeing itself as a set of processes. This made it easier to facilitate staff discussions around processes that needed to improve – in effect, the organization re-designed itself.

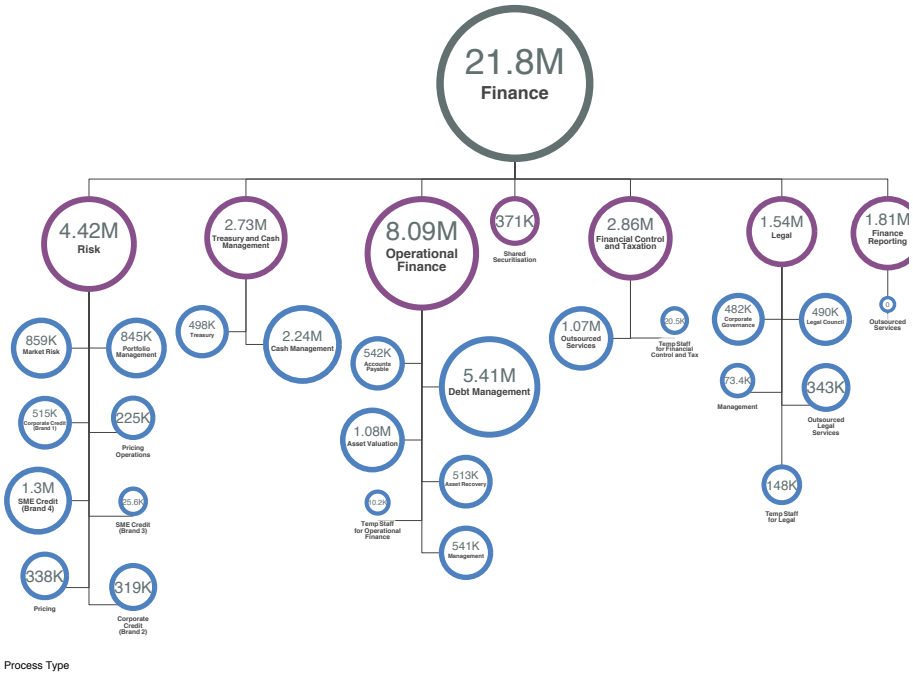

Fig. 3. Process Cost Visualization.
Source: Concentra

Incomplete data. Incomplete data might be seen as a problem for organization design. After all, an organization design is meant to treat the whole organization as a system – linking people to roles, to processes, to competencies, to client deliverables, and to objectives. Linked datasets are valuable in addressing incomplete data because they expose gaps. For example, a company may understand 100% of its costs. But in a changing world, can we link costs to processes? Can we link costs to clients? We have listed our organization's risks, but do we know who is responsible for handling each one? As people's roles change and outside factors alter risk levels, can we track who is overloaded? It is easier to sense-check this kind of analysis through linked datasets than it is through simple "one-aspect" datasets.

Data incompleteness also may be used deliberately. Google's 70-20-10 work system is based on the idea that the most valuable innovations may come from unexpected areas. Google has empowered its employees to spend up to 30% of their time on whatever seemed to them to be the most valuable use. It can be argued that this "unstructured" time is actually structured and managed very effectively. Peers review the work done, the choices made, and the results achieved in briefing sessions. Google's unstructured work time is an example of how organizations can be designed flexibly to include information gaps, to convert unstructured innovation into structured value.

Data reflexivity. Silverman Research's Social Media Garden allows a large group of people in an organization to consider ideas reflexively. Reflexive consideration means that people not only give their suggestions, but as Figure 4 shows, can view a bubble map of each other's suggestions – including size, color, and location indicating others' interest and agreement – and can respond. This design tool encourages ideas to develop over several rounds, allowing the socially constructed mass of ideas to influence its own evolution.

Fig. 4. Idea Mapping and Rating in Silverman Research's Social Media Garden.
Source: Silverman Research (www.silvermanresearch.com)

This methodology for gathering group ideas genuinely differs from surveying due to its looped nature and differs from a "town hall" meeting because of its greater potential scale and anonymity. It can be used as a step in the process of organization design, to surface issues and evaluate options.

Linked data. During the organization design process, we have found it critical to be able to link aspects of the organization to one another, so that impacts throughout the system can be understood properly. Linking is necessary because the many-to-many relationships between one aspect of the organization (e.g., people) and another aspect (e.g., responsibilities) are difficult to model and maintain in normal datasets. Linking is vital because it helps the organization to be conscious of where it has specified its activities, skills, deliverables, and risks, and where it still has gaps.

Figure 5 gives a conceptual model of an IT infrastructure supply company in which OrgVue was used to map (a) clients into client segments, (b) services to client segments, and (c) people to the services that they carry out for clients. Such a mapping allows designers to understand the connections between the client segments and the true underlying cost, either at the client level or the service level. This is vital for understanding the true cost to serve per client and redesigning the organization's structure and workflow.

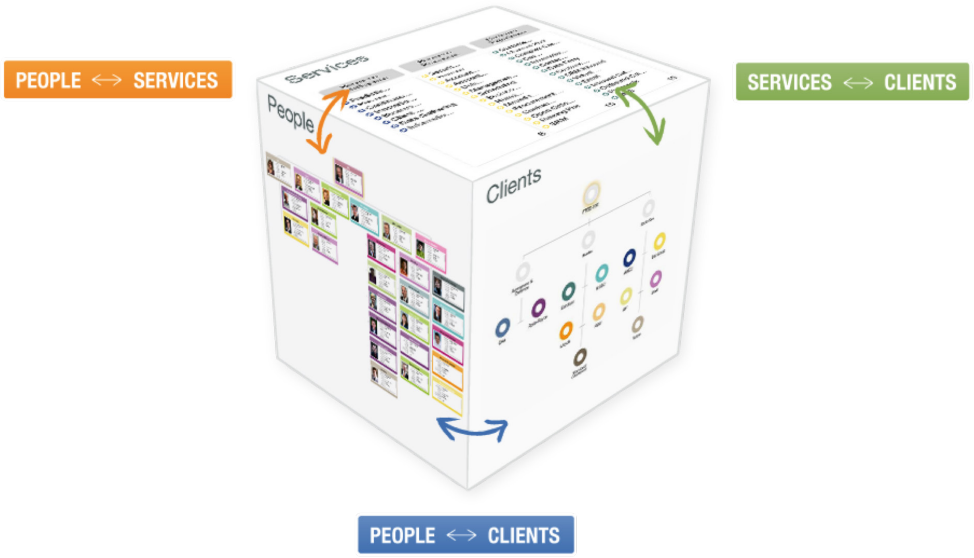

Fig. 5. Image Representing a Linked Model of People-Clients-Services.
Source: OrgVue (OrgVue.com)

Linking different aspects of people in an organization reflects the reality of organizational life. Linked connections, often mapped through graph databases (Webber, 2013), are fundamental to understanding organizations because:

- They are how humans work – but not how we are trained to think. We find it very hard to think in two connected dimensions at once, so we need systems that will let us agree on actions in one dimension and see their impact on other dimensions.
- They reflect reality. People may carry out multiple roles, have multiple skills, and deal with multiple customers or multiple products.
- They deal with the connectedness of change. When change occurs, organizations have to adapt as elements that are linked together. And those connected elements end up with a clustered, connected item: people.

CONCLUSION

The structures of organizations will certainly be different because of big data. We prefer goods that arrive on time, services on which we can give feedback, and recommendations that are tuned into our wants and needs. Big data can help with all of these desires. But the process of organization design is not a big data problem. The process of organization design is fundamentally driven by the bundled, reflexive, and linked nature of people data. People data are multiple-aspect with many-to-many links. Successful organization design in the future will make use of all the traditional tools, but it can avoid having to build enormous data warehouses. Instead, it will supplement the existing databases with graphing, visualization, and linking tools and methods that at last will let us treat organizations properly as systems.

Acknowledgements: With thanks for comments and feedback to: Rodin van der Hart, Marianna Favero, Dolly Mastrangelo, Richard Burton, Fabrizio Salvador, Tom Pape, Ben Marshall, Naomi Stanford, Will Sheldon, and others via the OrgVue blog site. Errors and omissions remain the responsibility of the authors.

REFERENCES

Allen K. 2010. Companies that put employee engagement into practice. *Observer* 22 August 2010.

Cross R. 2009. *Driving Results through Social Networks: How Top Organizations Leverage Networks for Performance and Growth*. Jossey-Bass, San Francisco, CA.

Davenport TH, Patil DJ. 2012. Data scientist: The sexiest job of the 21st century. *https://hbr.org.*

Farmer N. 2008. *The Invisible Organization: How Informal Networks Can Lead Organizational Change*. Ashgate Publishing Company, Burlington, VT.

Farmer N. 2013. E-mail to the authors, Informal Networks Ltd.

Galbraith JR. 2012. The evolution of enterprise organization designs. *Journal of Organization Design* 1(2): 1–13.

Galbraith JR. 2014. *Designing Organizations: Strategy, Structure, and Process at the Business Unit and Enterprise Levels*. Jossey-Bass, San Francisco, CA.

Laney D. 2001. Gartner definition of big data. https://www.gartner.com (Accessed on October 17, 2014)

McAfee A, Brynjolfsson E. 2012. Big data: The management revolution. *https://hbr.org.*

Silverman M. 2012. Social Media Garden, *Silverman Research Online*.

Slinger G. 2014. Dataset size for organization design. blog.orgvue.com, Concentra, February 2014.

Starr-McCluer M. 2000. The effects of weather on retail sales. Working Paper. Federal Reserve Board of Governors, Washington, DC.

Stephenson K, Lewin D. 1996. Managing workforce diversity: Macro and micro level HR implications of network analysis. *International Journal of Manpower* 17(4/5): 168–196.

Transport Catapult. 2014. *https://ts.catapult.org.uk/* (Accessed on February 8, 2014).

Webber J. 2013. A little graph theory for the busy developer. *SkillsMatter.com*, May 29.

GILES SLINGER

Product Management Director
Concentra
E-mail: giles.slinger@concentra.co.uk

RUPERT MORRISON

Managing Director
Concentra
E-mail: rupert.morrison@concentra.co.uk