6 Conclusions

This dissertation has examined the routes, processes and environments of knowledge production and transfer. At first glance, the overall route from knowledge production in a lab setting to the social interactions amongst firm founders located within a Science Park may appear convoluted.

The primary question of this dissertation (what knowledge elements are transferred from academia to industry, how are they transferred, and what factors influence this transfer?) draws upon larger theories of knowledge transfer in general and the effect of the environment, be it social or physical, on innovation. In order to answer the research question, we have to identify the researchers performing the process, the knowledge elements involved and the conditions under which knowledge transfer operates. This leads to the following sub-questions: (i) how can we disambiguate researchers with an effective balance between precision and recall?; (ii) how can we identify knowledge elements and their attributes in an operational way, and what elements are transferred between actors?; and (iii) what resources, and from which actors and operational spheres, contribute most significantly to the development of an academic spin-off and its host technology?

In this final chapter, I synthesise the conclusions of the four studies and how they relate to the primary question and sub-questions. These findings add to the existing knowledge and theoretical bases underpinning my research. The theoretical implications of the results and conclusions will be examined, with special regard to further applications.

6.1 How can we disambiguate researchers with an effective balance between precision and recall?

Scientometric studies have become increasingly important in the evaluation and analysis of research systems. Key to these evaluations and analyses is the availability of clean input data, and efficient and balanced disambiguation techniques are vital. We developed a method that relies on using the best possible combination of metadata available to us, based on the pitfalls of current techniques.

Three problems are commonly found with current techniques: discarding of data, limited selection of metadata, and a lack of consideration for evolving research streams and actual contributions to research by authors. Our solutions to handling the first two problems, data discarding and limited metadata, are interrelated: we selected the best possible alternate combination of available metadata. In other words, for records missing one or more fields, the proportional discriminating powers of the available metadata were adjusted depending on available combinations. All this occurs ‘on the fly’ within the algorithm, based on previous calibrations. The last problems, those of the evolution of terminologies (Healey et al., 1986), research streams or knowledge homogeneity (Tang & Walsh, 2010) and author contributions (Bates et al., 2004; Moed, 2000; Yank & Rennie, 1999), rely on recognising the changing roles, topics and requirements of researchers as they progress through their academic life cycle. For department heads, author status is occasionally ‘honorary’, in that much of the research and write-up was actually performed by others. In evaluations, however, every author listing counts. In other current techniques, similarity calculations assume input from all authors, in equal parts. Relying on this
assumption leads to networks of publications by authors, exhibiting low or no similarity. The algorithms developed in Chapter 2 tackle this problem by assuming that contributions by authors vary according to listing (an exception is made for alphabetically ordered listings). The discriminating powers of the indicators were adjusted to allow for this. For example, the importance of the title or abstract words (most likely selected by the authors writing the paper) is reduced, whereas the journal name (an aspect more likely to be decided on strategic grounds (Leydesdorff et al., 1994) by the head of department) is granted more discriminatory power in the algorithm.

A similar practice is deployed to adjust for the evolution of research streams. Over time, a researcher’s usage of words will change to reflect changing fields of interest and publishing. When comparing records published ten years apart, by apportioning less discriminatory power to title or abstract words but more power to journal field characterisation or cited references, we account for the evolution of researchers’ interests. Taking the dynamic selection of alternate metadata together with the adjustments for evolving research interests and contributions means that the developed algorithm allows for more accurate and inclusive results. The work conducted in Chapter 2 both provides theoretical insight into the research and publishing practices of individuals and also contributes to the larger issue of disambiguation.

6.2 How can we identify knowledge elements and their attributes in an operational way, and what elements are transferred between actors?

In the third and fourth chapters of this thesis, we developed a novel method for the visualization and quantification of specific contributions, by individuals and institutions, to the developmental routes of specific quanta of knowledge. We then applied this method to an individual firm founder, looking at (a) how the scientific background of the patent corpus links to the scientific output of the inventor; (b) how a researcher operates in a collaborative environment, and if the contributions of those contributors are visible in the patent corpus; and (c) whether the inventor demonstrates a level of adaptive knowledge acquisition, necessary for the development of a technology.

To link patent and publication data, we used the non-patent literature references (NPLR) found in patent applications. We grouped together the NPLRs and the author/inventor’s corpora of publications, linking their shared title words and cited references, and then they were clustered. The topical foreground and cognitive background of the research cited by the patent applications, and that of the author/inventor’s whole corpus of publications, provided a clear view of the required knowledge platforms and absorptive capacity of the inventor and of the contribution it makes to the development and transfer of knowledge. Both the bibliographic (examiner) and in-text (applicant) NPLRs were included to provide better balance between what is considered important to the technologies (Karki, 1997; Tamada et al., 2006), in order to negate some of the effects of strategic citing in patent applications.

For further clarification, each of the inventor/author’s research streams was differentiated further by the introduction of ‘concept clusters’. With these concept clusters, we identified the specific, rather than general, contributions made by the researcher involved. Our approach took into account the role of co-authors and co-inventors, allowing us to determine the specific expertise these collaborators added to the technologies. The institutional affiliation of the inventors and authors also gave an institutional approach to the contributions to the technologies. Additionally,
the multidirectional aspect of knowledge and skill transfer between basic and applied research can be examined in minute detail using this method, including the absorptive capacity of individuals and institutions.

In relation to the research question of this dissertation, it is important to remember that technologies should be seen in the context of their surroundings, practices, artefacts and base understanding (Nelson, 2004). Absorptive capacity (Cohen & Levinthal, 1990) and the theoretical extensions of acquisition, assimilation, transformation and exploitation under potential and realised absorptive capacity (Zahra & George, 2002) are necessary to describe the roles of endogenous and exogenous sources of knowledge in the development of a technology. The transfer of knowledge between universities and start-ups is most often based upon a specific technology or scientific result (Baba et al., 2009; Carayannis et al., 1998). This can occur in various forms, such as technology or skills (Steffensen et al., 2000) or collaborations (Agrawal et al., 2006). They are typically either codified, such as publications and patents, or tacit, notably skill sets (Cohen et al., 2002). The role of a star or core scientist (Furukawa & Goto, 2006; Zucker & Darby, 1996) is important in facilitating this transfer.

To aid the analysis of specific contributions by individuals to a technology, several descriptors were introduced for (1) the reputational and applicability aspects of the scientific base work conducted by an individual; (2) the localisation of an individual’s overall research trajectory in or outside the field(s) of research necessary for the technologies; (3) other fields of science being utilised by the technologies; (4) the level of input of collaborators; (5) the shared knowledge features (such as concepts, knowledge bases and, to a certain extent, skill sets) utilised by the technologies in relation to their sources; and (6) the individual incorporated skill sets acquired during the development of the technologies, and possibly applied to further basic research (knowledge creation feedback).

We found that during the initial stages of a technology’s development, our individual in the case study, Professor Yusuke Nakamura, recognised the importance of exogenous knowledge sources. The expertise of his network of co-inventors and co-authors increased his scope and ability to source new knowledge required for the technologies. The technologies were initially based upon Nakamura’s knowledge and skills set, but some aspects were outside his expertise. To combat this perceived gap in Nakamura’s base knowledge, the assimilation of new knowledge was necessary, and this was ultimately visible in the patent and publication analysis output. Research cited by patent applications that did not fall within Nakamura’s or his co-inventors’ expertise was quickly incorporated into the research agenda. The result of this research was seen in the increased publication output of Nakamura in these sub-fields and in some cases the eventual citing of this supplementary research by the patent applications in subsequent developmental stages of the technologies.

### 6.3 What resources, and from which actors and operational spheres, contribute most significantly to the development of an academic spin-off and its host technology?

In the fifth chapter, we examined, in tandem, the social and scientific networks of the founders of biotechnology-orientated firms in the specific context of a Science Park. We were looking to
investigate the cognitive development routes of an idea generated in academia and exploited in industry, the relations supporting the knowledge capture and transformation, in particular the role and sources of the firm founders’ social capital. Finally, we investigated the role of Science Parks in facilitating these processes.

Science Parks, and their utility, have been extensively studied, yet common definitions are hard to come by. General descriptions of a Science Park sum to a property-based, technology-orientated agglomeration of firms of varying specialisation and size, with close links and opportunities, either cognitive, geographical, structural or commercial, amongst firms and to a higher education or research institution (Das & Teng, 1997; Löfsten & Lindelöf, 2005; Quintas et al., 1992; Siegel et al., 2003). Each Science Park has unique origins - some were developed for the infrastructure, whereas other were developed to improve R&D innovation and production, or to provide intellectual development (Koh et al., 2005).

For firms choosing to locate within a Science Park, neoclassical location theory tends to dominate the decision processes of the firm founder (Westhead & Batstone, 1998). These typically include logistical issues, such as the proximity to the founder’s home. This is not to say that firms interviewed considered only these issues, but rather that practicalities won over potentialities. In our interviews with firm founders, they all expressed interest in being able to collaborate with other firms in the Science Park (in line with one of the espoused benefits of locating within a Science Park). However, there was little evidence in the publication and patent data to show that they actually conducted collaborative research with other firms located at the Park.

That is not to say that there was no collaboration, rather there was no substantial evidence of collaboration. From the patent and publication data, the regional and international characteristics of co-assignees and co-authors show that for almost all of the firms, collaborative activities were common, but with firms outside the Science Park. Academic collaborations were primarily with the local HEI, Leiden University in this case, and a few founders maintained strong links with their alma maters beyond Leiden. This was reflected in the interview data where the interactions were internal, i.e. initiated before firm’s formation and location to the Science Park, and external, i.e. with academic and industrial partners elsewhere in the country and abroad.

Social capital as a resource can be considered supplementary and enabling to the stock knowledge, financial capital and skills of an entrepreneur (Dubine & Aldrich, 1991; Greve & Salaff, 2001; Lin, 1999). In the interview data, for the firm founders who exploited their networks, in most of the spheres discussed in the chapter - particularly the scientific, technical and financial spheres - drew capital from either internal sources (i.e. historically through personal relationships prior to firm formation) or from sources external to the Science Park. Only a few firms reported any interactions of any nature with the Science Park administration or other firms located at the Science Park. Developing social capital through the discovery of opportunities, securing resources, and obtaining legitimacy (Elfring & Hulsink, 2003) means that the sources for these processes did not come from within the Science Park.

We found that the scientific capabilities of the firm founder were significant in developing and expanding the firm’s scientific base, and for its eventual patent output. The substantial similarities
between the patent content and the scholarly output of the firm founder (our proxy being the co-location of NPLRs and the founder's publication corpora) and the number of active research streams at and after incorporation both suggest that the scientific base of the founder had a large supporting role.

Debates surrounding the utility of Science Parks are bound to continue. Future research is likely to feature arguments relating to what specific market failures a Science Park addresses (Siegel et al., 2003), with the supplied rationales ultimately undermined by inconsistencies in how Science Parks are defined. Their broad descriptions (examples of which are presented earlier in this section) encompass Science Parks of a range of sizes, with varying intensity of ties to HEIs, and with different administrative styles etc. Studies comparing Parks can, unfortunately, only be applicable to the Parks mentioned in their data sets. Additionally, each Science Park has unique origins and unique motivations for its formation. These again relate to the specific market failures being addressed.

The lifeblood of a Science Park is its tenants. Science Parks compete for tenants, and those tenants exist in a coopetitive environment (Nalebuff & Brandenburger, 1996). Tenants compete for access to networks and the benefits these networks bring. They also are presented with the opportunity to cooperate with other tenants in the park, sharing resources and mitigating risks whilst expanding their potential access to networks. To counteract the diversity issues in a Science Park in terms of evaluating their utility, we feel that more emphasis is needed on analysing the knowledge structures and competences of the firms within a Science Park.

6.4 All together
The overarching research question of this thesis (what knowledge elements are transferred from academia to industry, how are they transferred, and what factors influence this transfer?) is at first glance a broad question. It is necessarily broad so as to encompass the complexity of knowledge transfer in relation to absorptive capacity, social capital and the environment in which these knowledge transfer processes take place. In anticipation of the inevitable question of applicability, this thesis goes some way to providing a toolbox for parties that are interested in discovering which elements are transferred, where to and where from, and what factors influence this transfer, for their specific field of application. Each chapter in this thesis provides a methodological approach that can be applied to different cases, in different contexts. The case studies in this thesis are used to provide examples of how the methodologies should be applied, but also to validate their logic and results. As such, each chapter provides substantive examples of each element of the primary question with its own case study.

The underpinnings of this methodological toolbox begin with considering the effect of an individual’s previous research on future research plans, and the similarities between past and present current research streams. The methods developed and insights provided in the chapter on disambiguation allow us to analyse factors such as the similarities and differences between research conducted during the PhD phase of a researcher’s career and the professorial phase. Over time, an individual’s research contributions may change with rank and eventual specialisation, but their incorporated knowledge and skill sets developed remain. Research conducted in academia and eventually applied in industry follows a convoluted path. We needed to gain an
understanding of this path for our disambiguation algorithms to succeed, and such an understanding provides a first glimpse at what knowledge elements are transferred over time.

The methodology and case study in the third and fourth chapters serve as a vehicle to examine the specific contributions of an existing knowledge base to the development of a technology platform, which involves identifying the knowledge elements and, to a certain extent, how they are transferred. The knowledge base does not necessarily come from one individual, but also from co-authors and co-inventors, and from other researchers working in different research settings. The methodology outlined in these chapters provided a toolkit for us to uncover the linkages between research conducted within academia and the eventual application of that research in industry, and the case study provided an example of what our approach can reveal. By applying our new method to a real case study, we demonstrated the ability to combine exogenously generated knowledge with a current knowledge base. New research conducted by Nakamura was guided by previous efforts, indicating that research practices and results are constantly evolving to inform, guide and provide the basis for extensions to different technologies. With detailed descriptions of the linked chains of research, we showed that a research corpus of an individual and their co-inventors and co-authors, can be readily recognised and identified in the exploitation of their research (i.e. in patents). The thematic links between the technologies and the underlying science was clearly identified in this chapter, verifying our methodological approach.

There is a long list of purported benefits for a firm to decide to locate to a Science Park. Its close proximity to a HEI and the prevalence of like-minded firms in the vicinity are mere examples of the reasons firms decide to locate to a Science Park. In social capital terms, a Science Park provides accessibility to resources, but for the firms in our case study, for the most part, these resources existed in potentia. It is important to note that for the firms under study there were no access barriers imposed by the Science Park. All the firm founders considered the Science Park to be an important potential source of collaborations and customers. If the opportunity arose, all the firm founders stated they would consider it. However, using the lens of only the purported benefits of locating to a Science Park, we failed to see more than practical benefits.

It is the firms that drive the success of a Science Park, and each firm is driven by its own scientific capacities and potential market linkages. For firms to truly enjoy the network benefits of a Science Park, there should be an overlap of not only their fundamental or applied sciences, but also of potential collaborators and customers. As such, we feel there needs to be a greater emphasis on the underlying sciences and technologies hosted by each firm when setting up a Science Park.

An important conclusion from this chapter was to view the primary research question from two perspectives: access to resources and technology development. We found it was necessary to blend the two perspectives, as many of the opportunities in one arise from the other. Ideas generated by a scientist in academia are not initially beholden to entrepreneurial dynamics within their network. They are, however, subject to the incentives structure of academia. New research streams are common in academia, where networks of scientists contribute to one another’s research, iteratively guiding the development of an idea or technology. If a certain stream or idea
is deemed suitable for exploitation, the scientist/entrepreneur/founder's various networks simultaneously come into play, and the opportunities for further scientific development diminish. In their place, commercialisation of the idea becomes paramount.

Incentive structures shape both the strategies of academic researchers and industrial researchers in terms of valuing their research and results, and thus what aspects of the research will be developed and transferred. For academic spin-offs, access to scientific and technical networks and/or resources is not restricted to academia but extend into industry. The development of these resources is crucial to the development of the technology, but is secondary for the spin-off itself. It is access to the organisational, financial and regulatory networks that tops the priority list and begins to refine the technology.

Referring back to the quote by Leonardo Da Vinci at the beginning of this thesis, “the colour of the object illuminated partakes of the colour of that which illuminates it”, this message becomes clearer when applied to knowledge transfer. The characteristics of an idea conceived in academia and transferred to industry vary in many ways but, using the developed tools, we can identify its origins, track its evolutionary path, and examine the effects of the environment.

6.5 Implications

From Chapter 2, in regards to future disambiguation, our approach confirms the need for an understanding of scientific research practices. Methods that employ purely statistical approaches often fail to accommodate the vagaries of research in practice. Each field has its own practices, which need to be taken into account in any disambiguation approach. For example, low citation rates, or a greater prevalence for single-author publications: both of these will affect an algorithm's discerning power.

A reliable and efficient disambiguation approach is extremely relevant to policy formation. For evaluation purposes, if corpora of publications are incomplete due to inaccurate disambiguation techniques, it can result in tremendous potential losses of funding for researchers. Furthermore, the rise of science systems from Asia, where there is a lesser degree of diversity amongst names, will make the issue of cleanly disambiguated data even more pressing.

There are other fields which have applications for disambiguation techniques, such as in social networks analysis and in search engine design. A future extension of our approach would incorporate heterogeneous data sets including professional social network data, and blog or self-published data, in addition to publication and patent data. The use of these ‘alt-metrics’ is growing in popularity as research - and science in general - is conducted at multiple levels of engagement and dissemination. The need for effective disambiguation techniques to cover the growing heterogeneity of data sources is absolutely required.

From Chapters 3 and 4, we believe that this method of mapping science to technology could deepen our understanding of the contributions made not only by individuals, but also by university departments and firms - providing us with data which can be used as input for theoretical systemic models such as those that concern funding instruments and policy. Technology positioning and evaluation models for funding allocation can be tested more completely by
examining the relative contributions of exogenous or endogenous sources of knowledge per field, allowing models to be adapted to suit varying publication and patenting output between fields. Policy-makers can be better informed on the adoption rates of indigenously-produced knowledge, with a deeper understanding of research competences in their own countries and how they compare to science systems elsewhere in the world.

In addition, the specificity of this approach is useful to all those involved in research. For universities, understanding specific contributions to a group of technologies can help TTOs to recognise what forthcoming research may be of value, either to their own IP portfolios or those of their industrial partners. This can similarly be useful on a regional or national level, with funding instruments gaining the ability to link and identify their monies or subsidies to specific topical areas in both science and technology output avenues.

Our method can be used in private R&D settings, with firms having a greater understanding of what potential avenues of research their in-house competences allow for, versus out-sourcing specific research-intensive tasks instead. Venture capital firms would find this method useful when determining the suitability and sustainability of the knowledge and knowledge producers involved in potential ventures. In short, this method enables a fine-grained approach to determining the applicability of past research to a future technology, which was previously not possible.

The implications of the research conducted in this thesis focus on (i) the methodologies for identifying and tracking knowledge transfer, and (ii) conditions for knowledge transfer in general and, more specifically, the Science Park as both a driver and an environment for knowledge transfer. The analysis of Chapter 5 leads to several implications for Science Parks. Science Parks represent a massive investment in financial terms, which of course will factor into technology development and knowledge and technology transfer for many, if not most, academic spin-offs. In scientific literature and policy discussions, the subject of Science Parks has typically being marked by conflicting opinions and findings. Studies routinely either conclude that Science Parks stimulate regional development or, alternatively, show no evidence of contributing to the local economy or innovative capacities of firms in them. These conflicting results stem from a lack of a consistent framework with which to evaluate Science Parks. Science Parks, in theory, should be enormous drivers of development and innovation. They provide the benefits of agglomeration, they provide a contact space for firms to interact, and they provide close proximity to a university and an educated workforce. But all too often, reasons cited by subjects in this study (and many others) for locating in Science Parks only relate to agglomeration and practical effects. Science Parks have a role in ensuring that there are enough ‘different but similar’ tenants to not only encourage the growth of firms in the park, but also the Science Park itself, and the many industrial partners and universities and government looking to see their investments being capitalised on. Each Science Park can, and should, provide an ecology of firms and knowledge that is autocatalytic in nature, providing inputs and outputs that can be utilised by most, if not all, who are located there.
6.6 References


