

Knowledge in motion: the evolution of HIV/AIDS research

Ryan Light¹ · jimi adams²

Received: 19 September 2015
© Akadémiai Kiadó, Budapest, Hungary 2016

Abstract Many contemporary social and public health problems do not fit neatly into the research fields typically found in universities. With this in mind, researchers and funding agencies have devoted increasing attention to projects that span multiple disciplines. However, comparatively little attention has been paid to how these projects evolve over time. This relative neglect is in part attributable to a lack of theory on the dynamic nature of such projects. In this paper, we describe how research programs can move through various states of integration including disciplinarity, multidisciplinarity, interdisciplinarity and transdisciplinarity. We link this insight to computational techniques—topic models—to explore one of the most vibrant and pressing contemporary research areas—research on HIV/AIDS. Topic models of over 9000 abstracts from two prominent journals illustrate how research on HIV/AIDS has evolved from a high to a lower level of integration. The topic models motivate a more detailed historical analysis of HIV/AIDS research and, together, they highlight the dynamic nature of knowledge production. We conclude by discussing the role of computational social science in dynamic models of interdisciplinarity.

Keywords Interdisciplinarity · HIV/AIDS research · Topic models · Dynamic networks

Introduction

Attempts to solve the HIV/AIDS crisis have engaged a diverse range of topics from disease diffusion to vaccine development and draw on research from a wide range of academic disciplines from sociology to chemistry. Like other pressing social and public health issues

✉ Ryan Light
light@uoregon.edu

¹ Department of Sociology, University of Oregon, 1291 University of Oregon, Eugene, OR 97403-1291, USA

² University of Colorado Denver, Denver, CO, USA

spanning multiple disciplines, the extent of overlap between the research fields investigating HIV/AIDS is unclear. The prevailing wisdom is that interdisciplinary work increases the development of “new kinds of knowledge” (Boyack et al. 2005: 372; Jacobs and Frickel 2009: 48): Social network analysis has long shown that there are net gains to drawing from a diverse range of resources and academic research may follow course (Burt 2004; Powell et al. 2005). Yet, despite the resources and attention dedicated to it, interdisciplinarity remains under-examined (Jacobs and Frickel 2009). For example, many research and funding strategies assume that interdisciplinarity is static. Recent research points to the dynamic quality of knowledge production (Jacobs 2013; Leydesdorff and Schank 2008), but how these processes unfold, and what implications that has deserve further attention. Building a more complete, dynamic understanding of interdisciplinarity is essential as many contemporary, social problems, such as the HIV/AIDS epidemic, likely require some balance between specialization and the cross-pollination of ideas.

With this in mind, we analyze how HIV/AIDS research has changed over time. HIV/AIDS research is an interesting case of knowledge production because it has evolved relatively quickly and experienced several identifiable successes, such as disease identification and treatment. Yet, at the same time, understanding knowledge production within the context of HIV/AIDS research remains important as the community has to date failed its ultimate objective: Developing a sure-fire means of stopping the spread of HIV/AIDS. This case also illustrates more general lessons about research on contemporary social and health problems best conceived as “in motion.” In other words, by elaborating a dynamic picture of knowledge production, our approach expands earlier typologies of interdisciplinarity by focusing on the content of a research area as it changes over time. Some research problems, often those dealing with the most pressing social and/or medical issues, demand some type of integration across disciplinary boundaries. Yet, this integration is neither simple, nor static. Indeed, the content of research is subject to dramatic changes as practitioners promote or resist moves towards interdisciplinarity.

In this way, our dynamic approach to interdisciplinarity conceives of research as, in part, an example of boundary-work (Crane 1972; Gieryn 1999). In this framework researchers construct, maintain, and challenge the boundaries between disciplines when facing new research problems (Jacobs 2013; Knorr-Cetina 1999). Boundary-work is inherently dynamic, and as such, adds the element of change to prior, and more static, conceptions of interdisciplinarity. Our approach also suggests that the evaluation of a project is better understood through these dynamic trajectories than by identifying states at any singular time in that project’s evolution. Overemphasizing particular moments in the evolution of a research project could lead to a biased view of its integration and influence evaluations of a project’s success. To trace how the content of HIV/AIDS shifts over time, we use topic modeling, a technique developed in the computer and information sciences to uncover latent themes within a collection of texts, such as scientific abstracts (Blei 2012; Mohr and Bogdanov 2013; Steyvers and Griffiths 2007).

Using topic models, we examine the history of HIV/AIDS research through a macroscopic lens uncovering the broad shifts in its content. We map the dynamic topic structure of 9067 articles published in the prominent journals *AIDS* and *JAIDS*. This model allows us to identify the emergent characteristics of HIV/AIDS research as uncovered through a global view of the content of the research itself. A socio-historical analysis structured by a longitudinal series of topic model networks indicates that HIV/AIDS research began as a robust, interdisciplinary project, but has grown less so over time as specialists have carved out particular niches along a behavioral/clinical/bench science divide. Despite being both a social and biomedical problem, many topics within attempts to address the HIV/AIDS

crisis have become more entrenched in specific disciplines. We conclude by discussing the role of computational social science in dynamic models of interdisciplinarity.

Boundary-work and many disciplined science

While the notion of “interdisciplinary” research has captured recent attention, the act of scholars crossing disciplinary boundaries is not new. Boundary-work—or attempts to maintain and transgress boundaries—plays an important role in the history of disciplinary development. The foundation of sociology, for example, is at its core the cumulative boundary-work of 19th and early 20th century scholars, who borrowed concepts from biology, ecology and other disciplines (Salter and Hearn 1997). The history of sociology is one case among many: knowledge production has often involved disciplinary cross-pollination and even theft. Yet, interdisciplinarity becomes more visible by the end of the twentieth century as states and institutions dedicate substantial resources towards the promotion of interdisciplinary engagement. Surprisingly, the value of the interdisciplinarity has gone largely unchallenged with relatively modest research critically examining interdisciplinary projects (see Frickel et al. 2016; Jacobs and Frickel 2009; Jacobs 2013 for exceptions).

Valuable research on knowledge production constructs important typologies outlining the boundaries between various strategies for organizing and producing knowledge. Yet, this prior work generally overlooks the dynamic quality of knowledge production. Here, we build upon prior definitions of interdisciplinarity. We also add to this a picture of interdisciplinarity that builds from research on boundary-work to emphasize how specific states of knowledge production, such as interdisciplinarity or multidisciplinary, are subject to change. Unfortunately, “interdisciplinary” is often used as the catchall term for all of the potential states of knowledge production (Klein 2010; NAS et al. 2005). For this reason, here we rely on the term “many-disciplined science” to incorporate the undifferentiated collection of research approaches bridging multiple disciplines. Within this literature, there are at least three forms of many-disciplined science: cross-/multi-disciplinary, interdisciplinary, and transdisciplinary (NAS et al. 2005).

The most basic level at which many disciplines can engage a research topic is what is typically known as multidisciplinary, or cross-disciplinary, research. Multidisciplinary research engages any single research area from the vantage point of multiple disciplines, but does so in a way that maintains each discipline’s traditional boundaries, minimizing cross-fertilization (Klein 2010; NAS et al. 2005). This approach poses little challenge to the specialization that arises from more disciplinary-based approaches. Medical teams often take multidisciplinary approaches to care. For example, tumor boards use a multidisciplinary strategy to construct comprehensive treatment strategies for cancer patients. These boards consist of surgeons, oncologists, radiologists, and others who conference with one another to develop treatment plans for their patients, but remain firmly situated in their respective specialty areas (Newman et al. 2006; Westin and Stalfors 2008).

Following the National Academy of Sciences, interdisciplinary research is a more integrative form: Interdisciplinary research consists of research conducted by teams or individuals that “integrates information, data, techniques, tools, perspectives, concepts, and/or theories from two or more disciplines or bodies of specialized knowledge to advance fundamental understanding or to solve problems whose solutions are beyond the scope of a single discipline or field of research practice” (2005: 26). This definition highlights the bridging of disciplinary boundaries as something that happens at the level of

individual—or teams of—scientists—i.e., by research *practitioners* whether individually or collectively. Such efforts are made to draw on the variable strengths of different disciplines, and are brought to bear on advancing developments to the project that an individual/team is aiming to contribute to (Gondal 2011). For example, interdisciplinary work by bioinformatics researchers has combined biology and computer science to make great strides in advancing medical research (Choi and Pak 2006). With the wealth of genomic data, in particular, the ability to combine computational techniques and biological expertise has revolutionized how medical researchers approach drug development (Bayat 2002).

While perhaps less visible than interdisciplinarity, transdisciplinarity is characterized by relatively complete elimination of the salience of disciplinary boundaries in the scientific enterprise. Within the literature conceptually clarifying the typology of many-disciplined science, transdisciplinarity's boundary-less science arises via two quite different processes. The first form comes from a reorganization of scientific efforts—starting with the identification of questions, and extending straight through the strategies employed in answering them—around problem-based areas. That is, research is no longer bound up within disciplinary confines. To truly explain phenomena of interest, science cannot leave the framing of questions, methods of analyses, nor interpretation of evidence to the standards of any one discipline, but must instead span across those boundaries—problems have no respect of disciplinary boundaries (Hadorn et al. 2008).

Transdisciplinarity also takes the form of the joint problem-solving of academic disciplines and practitioners to solve “real-world” problems (Klein 2001; Nicolescu 2002). In other words, proponents of this variant of transdisciplinarity seek to transcend not only the boundaries between disciplines, but also the boundaries between the researchers and the public (Nicolescu 2002; Hadorn et al. 2008). This form of transdisciplinarity often highlights the importance of participatory or action research that emphasizes collaboration between researchers and the public (Couch 2004). For example, initial attempts to construct a successful treatment of AIDS were hindered by the conservative norms of bench science and tacit or overt biases. Motivated by the devastation in their communities, AIDS activists began to see the value of building expertise in the medical science on AIDS (Fujimura and Chou 1994). While most medical researchers were skeptical of this intrusion by the lay public, eventually a partnership formed. These “citizen scientists” helped develop the experimental protocol that led to the mid-1990s discovery of the treatment regimen that has saved tens of thousands of lives (Epstein 1995, 1996). In the end, citizen scientists most likely return to their non-research lives and interaction between lay experts and scientists wanes.

A dynamic approach to knowledge production

Typologies of many-disciplined science have challenged oversimplified notions of interdisciplinarity by pointing to the variety of forms that this kind of knowledge work can assume and by identifying the differentiation between their typical characteristics (Jacobs 2013; Porter et al. 2007). Clearly defining the various potential states of knowledge production helps avoid placing undue credit on one form of knowledge production over another. For example, as Jacobs (2013) describes, numerous multidisciplinary projects—e.g., universities' “institute” building—have been quite successful and lack the integration and institutional commitment of interdisciplinary projects. However, these conceptual distinctions have occasionally been lost when employed in analytic work examining the

relative advantages of these various forms. Both proponents of interdisciplinarity and scholars who are critical of many-disciplined approaches sometimes ignore the importance of these distinctions. One way this happens is when researchers assume the forms are conceptually indistinguishable, e.g., when criticizing many-disciplined research as a fad-dish “spirit of the times” (Hackett and Rhoten 2009). Another is when the distinctions are obfuscated by operationalization choices that lumps them together analytically, for example when all many-disciplined research is placed under the interdisciplinary umbrella (Hackett and Rhoten 2009). A more dynamic approach to knowledge production that conceptualizes interdisciplinarity as a state that is subject to change and, importantly, exists relationally alongside other forms of knowledge production helps focus attention on how or whether a particular research effort is taking a specific form, such as when or whether a project is multidisciplinary or interdisciplinary.

While prior work discusses interdisciplinarity through a boundary-work perspective (Dubrow 2011; Jacobs 2013; Small 2010), we add to this prior work in an important way: previous work often assumes that this boundary-work occurs as a singular *linear* process moving from disciplinarity through multidisciplinary and interdisciplinarity toward transdisciplinarity. In this formulation each progression corresponds to a step-wise lowering of disciplinary boundaries. Our contention however is that boundary-work does not require or even frequently result in this idealized progression. Rather, knowledge production can assume numerous different sequences.

Since Gieryn (1983, 1999) the idea of boundary-work has focused on the effort of managing boundaries. Scientists engage other scientists to collectively answer a problem or, alternatively, to resist co-optation, pushing and pulling against the intersection of their areas of expertise. This effort can take place quietly as one scientist agrees to work with another, subtly within university offices as one department resists an interdisciplinary research center, or visibly in the pages of major newspapers as academics from one discipline seek to shame those in another. In any case a boundary-work perspective on interdisciplinarity requires a parallel acknowledgement of the socio-historical nature of academic labor.

One location of boundary work is the field structuring “discursive knowledge” produced by scientists (Lucio-Arias and Leydesdorff 2009). Recognizing the discursive property of knowledge production emphasizes both the textuality of science—in so much as scientific debate occurs via words and symbols—but also its temporality. Yet, the categorical nature of prior work relies on a priori assumptions of difference, which makes it difficult to evaluate the socio-historical characteristics of knowledge production. This static conceptualization hides the fact that intellectual projects change due to competition over funding, progress on the research front, changes to national funding priorities, and so forth. They come into being, they live, and they die or evolve into some more stable organization, such as a discipline (Bettencourt et al. 2008). To avoid this pitfall our approach identifies boundary-spanning as a dynamic process that occurs along continua. This dynamic model encourages less rigid categorization as knowledge production does not shift in stepwise fashion from one state to another. In other words, our framework allows us to elaborate more explicitly how research projects transition between the various identified forms and does not simply assume a linear increase/decrease in disciplinary integration.

Below we describe computational methods for identifying research topics within the HIV/AIDS research field. We then draw on three features of these identified topics as a means to evaluate how HIV/AIDS research changed over time. First, we examine the content of the identified topics, highlighting how readily those topics correspond to traditional disciplinary questions or emergent problem-oriented questions associated with the discovery of and engagement with a new disease. Second, we identify the temporal patterns in the relative

popularity of those topics. While some aspects of the field's trajectory are somewhat difficult to predict we do expect that topics consistent with the exploratory-side of research are likely to wane over time. For example, the more recent research is likely to move beyond the development of disease identification mechanisms as solutions to identification grow more settled. At the same time, research on treatment is likely to expand as research on HIV/AIDS moves beyond the initial identification phase. Third, we examine the "topic structure" of the corpus, which identifies the ways that topics within the corpus cluster together and/or develop relatively independently from one another. Consistent with our attention to the dynamics of knowledge production, we focus on how this topic structure evolves through time. We employ an emergent strategy for evaluating this topic structure to investigate the general pattern observed in the field: Does HIV/AIDS research begin from a less disciplinary position and move to a more interdisciplinary one? Does this trajectory stabilize or do divides persist throughout the twenty years examined in this analysis?

The emergent topic modeling reveals the locations of overlap within the discursive field structuring research on HIV/AIDS. Discursive fields reveal the outcomes of boundary work that takes place in research labs and offices that result in publication. Here, with evidence of the labor of science, we can evaluate the temporality of many-disciplined problems. To anticipate how we will interpret the overlaps within the topic structure as a means of evaluating the dynamic properties of knowledge projects, we combine (1) features of topic structure that are available from analyzing topic overlaps as networks, with (2) substantive interpretations of the content of those topics. If we identify limited overlaps between topics, that could arise via two contrasting processes. First isolated topics that align with traditional disciplinary questions would be indicative of low boundary permeability (i.e., disciplinary organization of the field). If instead topics are primarily identifiable as "problem" oriented, that would indicate high permeability of disciplinary boundaries; problem orientation would indicate a transdisciplinary approach drawing from multiple disciplinary perspectives to identify, examine and "resolve" those problems (e.g., reduce transmission, improve treatment, etc.).

In the primary patterns that will be demonstrated below, few of the topics observed will fit the level of isolation described above, instead revealing modest to high levels of overlap. This overlap will occur especially between topics that fit the classic substantive boundaries structuring HIV/AIDS research (e.g. the clinical, bench, and socio-behavioral sciences), but also between these common thematic areas. With groups of topics exhibiting relatively strong clustering between them—the ability to differentiate between more interdisciplinary and multidisciplinary organization again turns to how that observed clustering aligns with the substantive makeup of the topics that comprise the identifiable groups. A map of the topic structure with high overlap between disparate topics indicates an integration across disciplinary orientations (i.e., is interdisciplinary in nature). Alternatively, if projects span disciplinary boundaries, but the ideas used to solve them do not, the clustering that is identified reflects a similarity within cluster that is primarily driven by disciplinary boundaries (i.e., multidisciplinary).

Data

This project draws on 9192 articles published in the journals *AIDS* and *JAIDS*—which represent two premiere putatively interdisciplinary journals specific to HIV/AIDS research—from their respective inception through the end of 2010. The prominence of

these journals can be seen from their impact factors of 4.91 for *AIDS* and 4.57 for *JAIDS*, which rank them fifth and ninth respectively among 57 journals on infectious diseases (ISI 2012). Of those published articles, 9067 (98.6 %) include abstracts, the texts of which provide our primary data.

Combined these abstracts include 378,398 tokens, or words plus numbers. We subject this database to a series of standard reduction techniques that allow the analysis to focus on the meaningful language distinctions in the corpus (Blei 2012; Griffiths and Steyvers 2004; Grun and Hornik 2011; Steyvers and Griffiths 2007), which results in a working vocabulary of 3178 words. This reduction process began by removing delimiting characters, but maintaining hyphens and % signs as these may play a role in several conceptual tokens. Next, we ran the vocabulary against a standard stop list that included words that are so conceptually irrelevant to be meaningless (e.g., conjunctions, prepositions, articles). We also stemmed the list to the words' common roots—e.g., so that immunodeficient and immunodeficiency are identified as conveying practically the same information (Porter 1980). To create the vocabulary for our analysis, we also removed very common (e.g., HIV and AIDS, which occur in virtually all papers in the corpus) and rare words that are not likely be of “predictive importance” (Blei and McAuliffe 2010:13).

Methods

Computer and information scientists have developed robust strategies for analyzing large sets of unstructured text data. One promising line of such research is topic modeling, particularly latent Dirichlet allocation (LDA). Previous research on LDA topic modeling in science largely focuses on how research is organized to discover more efficient information retrieval methods (Blei et al. 2003; Blei 2012) or to advance our understanding of human cognition (Griffiths and Steyvers 2004; Steyvers and Griffiths 2007). In essence LDA is a statistical model of language, which attempts to “reverse engineer” the writing process (Mohr and Bogdanov 2013). The model assumes that texts consist of a distribution of themes or topics, which in-turn are produced from distributions of words. This hierarchical model—consisting of texts, topics and words—assumes that authors write texts by combining topics, and in writing about those topics they draw from the words that comprise them.

LDA estimation algorithms *simultaneously* attempt to identify how words are used, *in combination*, to construct topics, and how documents are constructed from combinations of those topics (DiMaggio et al. 2013). There are many “bag of words” approaches to identifying meaning from texts (Blei 2012; Griffiths and Steyvers 2004; Leydesdorff and Hellsten 2006). Where LDA fundamentally differs from other approaches is by focusing not just on distributions of words over corpora to identify topics, but doing so in a fundamentally relational way—by focusing on the co-occurrence of words in assessing their assignment to topics (DiMaggio et al. 2013). In doing so, LDA resolves the countervailing needs of allowing single words to probabilistically belong to multiple topics (e.g., in the case of HIV/AIDS research the need for “drug” to able to signify a risk factor in the case of intravenous drug use, and a treatment approach in the case of antiretroviral drug combinations), with allowing different words to indicate the same themes (e.g., as will be seen below, the theme of “sexual transmission” is marked by substantially different language in the early phases of the epidemic compared to more recent periods). An additional benefit of this approach is that documents do not have to be assigned to single

topics, rather they can be assigned with proportional probability to multiple topics (Blei et al. 2003).

Social scientists have recently begun to use topic models generally, and LDA in particular (McFarland et al. 2013; Ramage et al. 2009; adams and Light 2014), to examine a range of substantive questions including newspaper coverage of US governmental funding of the arts (DiMaggio et al. 2013) and differentiation between English and French streams of demographic research in the 20th century (Marshall 2013). Following Griffiths and Steyvers (2004) and Light (2014), we exploit the historical application of computational text analysis to evaluate knowledge production within the context of HIV/AIDS research, and, importantly, how HIV/AIDS research has changed over time.

Here, we rely on *lda* 1.3.2 in R (Chang 2012). The fact that topic modeling is unsupervised means that we do not need to know what core ideas connect the corpus (Ramage et al. 2009), which comes at a cost: how do we know how many topics structure the set? While still the subject of some debate, a common solution relies on perplexity scores which maximize the generalizability of topic models (Steyvers and Griffiths 2007) coupled with a project-based determination based on the research objectives (e.g. a 500 topic solution may be unwieldy for telling a sociohistorical story). Indeed, research by Chang et al. (2009) on exploratory uses of topic models, such as the sociohistorical use here, finds that human coders and traditional metrics often differ over whether topics are coherent: Human coders may locate meaning more parsimoniously. Chang et al. (2009) recommend optimization techniques that incorporate human judgement. While this intriguing direction is outside the bounds of this paper, we did evaluate both perplexity and human judgment, described next, when determining the 30-topic solution for this analysis.

Next, we attempt to identify the main characteristics of each topic. Characterizing and naming topics is often difficult and accomplished via an ad hoc process. Variations and advancements in LDA try to alleviate some of this concern, yet this remains an unresolved problem. Here, the probabilistic assignment of papers to each topic provides the primary analytic categories employed to identify the structure of the field. In addition the model output also provides a list of the top words associated with each topic. For our models, we provided the top-50 words associated with each topic to four separate HIV experts, whose collective training/research experience spans biology, demography, epidemiology, genetics, clinical medicine, sociology, vaccine development and virology.¹ Each independently named the topics and went through an iterative editing process that resulted in near unanimous characterizations. In the two cases where differences arose from the independent coding process, focused discussions helped to resolve conflicts.

The proof, as they say, is in the pudding and like most sociological research the creation and naming of groups requires care and cross-check, but rarely engenders the kind of certainty one might like. As such, the relative ease with which the topics were labeled by the four independent coders—after unsupervised identification—provides some face validity to the model. In addition to the analytic utility demonstrated in the findings and interpretations below, we suggest that additional confirmation and confidence of these models will arise through future additional usage, and through combining these methods with others with similar aims (e.g., manual codings of textual data via content analysis and/or grounded theoretic approaches) to validate and interpret the results.

¹ The top-50 tokens loading on each topic explain a significant amount of the variation between the topics and led to the relative ease our expert coders had labeling the topics. A smaller list may not provide enough information for reliable labeling, a longer list of tokens could introduce unnecessary noise into the labeling process.

We then construct networks of topics, which allow us to gain a clear picture of the structure of the content of HIV/AIDS research and how it changes over time. We build networks for four time slices by constructing correlation matrices for each period. Topics—the nodes or points—are closer together the extent to which they appear in common abstracts. The edges or ties are correlation scores. We retain edges with scores that are in the 90th percentile for each slice. To assist in evaluation of these topic networks, we examine the Louvain community structure (Blondel et al. 2008), which assists in identifying how those topics form overlapping groups (beyond the pairwise overlaps indicated by the edges). This resulting network can be seen as the backbone of research on HIV/AIDS for each time period. Importantly, our analysis of interdisciplinarity in HIV/AIDS research does not end, but rather begins with these topic networks: we use the networks as maps to construct a sociohistorical narrative about the evolution of HIV/AIDS research.

Results

Overview of topics

Table 1 presents an overview of the topic modeling results. These topics reflect relatively well-bounded topics as indicated by the sets of terms most associated with each. Several topics capture the efforts of the “bench” sciences, such as topics that pertain to animal models of AIDS progression (topic 2) and HIV-testing assay development (topic 25), while other topics address more social scientific concerns, such as ethno-racial/gender differences in HIV/AIDS prevalence, particularly in the US (topic 21), or socio/behavioral factors of HIV in Africa (topic 1). Several topics capture multiple dimensions of a more global topic. For example, a series of topics (topics 26–28) address various factors associated with transmission potential, including those associated with intravenous drug use (28) and sexual behaviors including “high risk behaviors” associated with commercial sex work and co-infection with other STDs (27) and differences between transmission in same sex partnerships compared to heterosexual relationships (28).

These topics obviously are not equally weighted across the corpus of articles. Figure 1 presents the topic frequencies within the dataset. The topic on cost effectiveness (7), capturing articles evaluating HIV mitigation program tradeoffs (especially comparing the returns between prevention and care efforts), appears most frequently within the corpus, while the hepatitis topic (11), articles addressing clinical interactions of HIV and viral hepatitis infections, is least common.

The wide range of topics provides initial evidence that the journals *JAIDS* and *AIDS* appear to capture the interests of multiple disciplines from genetics and microbiology to epidemiology and public health. If we stopped our analysis at the identification of the breadth of topics covered and disciplines involved, we may conclude that HIV/AIDS research is a robust, interdisciplinary project. Prior theory and static discussions of interdisciplinarity would lead us to this conclusion. However, a dynamic understanding of interdisciplinarity demands that we dig deeper into the life course of this research program: simply identifying how much ink is spilled on which topics insufficiently captures the history of HIV/AIDS research.

Table 1 Topic overview

	Label	Description	Top-5 tokens				
			1	2	3	4	5
1	Africa	Socio/behavioral factors of HIV in Africa	africa	rural	epidem	survei	adult
2	Animal models	Animal models of AIDS progression	immunodefici	anim	replic	macaqu	node
3	ARV1	ARV—metabolic complications	bodi	metabol	mass	lipid	proteas
4	ARV2	ARV adherence	medic	adher	score	self-report	care
5	ARV3	ARV toxicity in clinical trials	dose	random	toler	efficaci	safeti
6	Clinical trials	Anit-HIV drug testing	transcriptas	nucleosid	copies/ml	proteas	virolog
7	Cost effectiveness	Cost Effectiveness comparisons	cost	research	review	countri	strategi
8	Drug failures	Drug resistance/therapy failure	mutat	genotyp	phenotyp	gene	variant
9	Drug metabolism	HIV-drug metabolism	concentr	dose	pharmacokinet	curv	area
10	Genetic variability	Global virus genetic variability/evolution	subtyp	region	strain	phylogenet	genet
11	Hepatitis	HIV and viral hepatitis—Clinical interactions	hepat	hcv	co-infect	coinfect	liver
12	Immunity	HIV immunity (cellular)	cytokin	t cell	receptor	function	stimul
13	Methods1	Behavioral science quantitative methods	odd	regress	multivari	[ci]	logist
14	Methods2	Basic/bench science methods	predict	sensit	valid	base	calcul
15	Neuropathy	HIV-related neuropathy	fluid	brain	neurolog	system	nervou
16	Oncology	HIV oncologic complications	sarcoma	kaposi	lymphoma	cancer	lesion
17	Opportunistic infections	AIDS-related opportunistic infections	pneumonia	hospit	pneumocysti	carinii	tuberculosi
18	PCR	Polymerase chain reaction	chain	reaction	polymeras	pcr	mononuclear
19	PMTCT	Prevention of mother-to-child transmission	infant	mother	children	matern	pregnanc
20	Prevalence	HIV/AIDS prevalence estimation and surveillance	hospit	27 %	versu	12 %	17 %
21	Race/gender	Ethnoracial/gender differences	state	diagnos	trend	unit	white
22	Replication Blocking	Molecular blocking of HIV-replication	inhibit	protein	replic	line	cultur
23	Survival	Survival PLWHA	hazard	death	surviv	mortal	kaplan-mei
24	Symptampotology	HIV-specific morbidity	stage	marker	asymptomat	serum	immunolog

Table 1 continued

			Top-5 tokens				
Label		Description	1	2	3	4	5
25	Testing assays	HIV-testing assay development	seroposit	western	serolog	screen	blot
26	Transmission1	Sexual Transmission of HIV	condom	intercours	unprotect	survei	anal
27	Transmission2	“Risky” sexual transmission of HIV	transmit	genit	std	worker	condom
28	Transmission3	Intravenous drug use	user	idu	needl	share	syring
29	Treatment	Treatment efficacy	suppress	copies/ml	haart	virolog	interrupt
30	Vaccine	HIV vaccine strategies	vaccin	neutral	peptid	antigen	protein

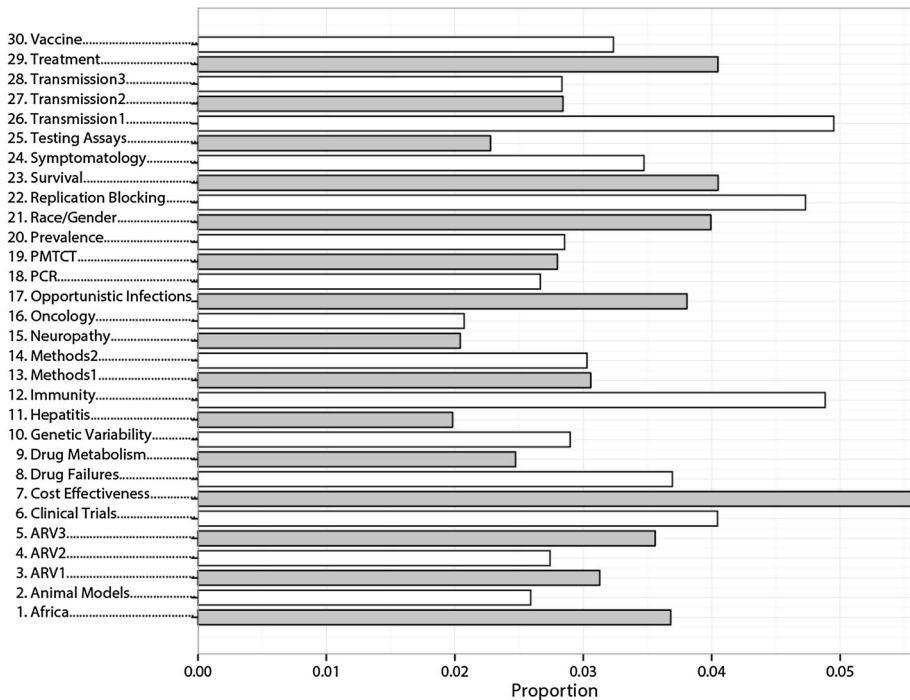


Fig. 1 HIV/AIDS research topic distribution

The rise and fall of topics

A dynamic approach to knowledge production suggests the high likelihood that topics will grow in and out of favor. This is particularly true with the problem-based orientation within HIV/AIDS research, involving some progression from identification to amelioration. If we have a stable means of identifying the existence of a problem, then the content of that research program will likely drift away from issues of identification over time. Figure 2 depicts the topics that have changed the most in the research covered in this analysis. The topics that experienced the greatest decline from 1990 to 2010 (Fig. 2a) largely capture content that pertains to the development of testing or identifying the disease. For example, symptomatology (topic 24)—the study of the series of symptoms associated with a disease—has experienced the most significant decline. Along similar lines the development of testing assays (topic 25) has also experienced steep declines. The development of widely useable tests is one of the relatively “settled” questions in HIV/AIDS research: Scientists know how to identify the virus and when people have it with relative ease (McKenna et al. 1997). Similarly, topics on opportunistic infections (17) and blocking viral replication (22) were fundamental questions in developing an understanding of how HIV/AIDS progressed among people living with HIV/AIDS.

Figure 2b depicts the other half of the equation: While some topics go into decline, others are likely to rise and fill their void. Within the context of HIV/AIDS research, the initial focus on issues of testing and identification declines, while topics that perhaps indicate a more mature science become more popular. For example, cost effectiveness

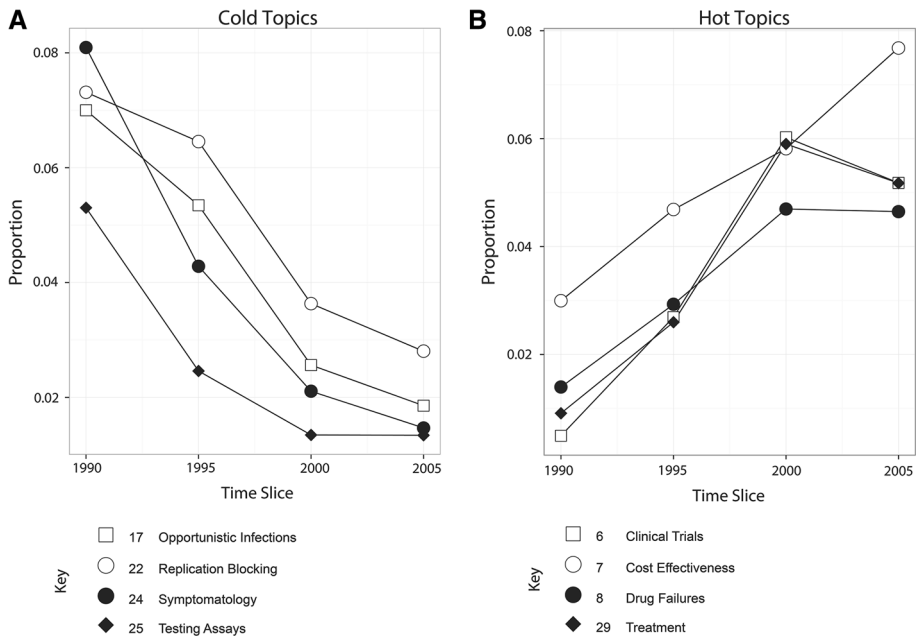


Fig. 2 HIV/AIDS research topic trajectories

(topic 7) experienced the greatest increase within the corpus. This is consistent with competing efforts between research teams and funding agencies to address a variety of concerns from disease prevention to care. Similarly, the rise in interest of treatment (topic 29) signals the growing attention to care regimens in impeding the disease; however, the steady rise of drug failures (topic 8) indicates the limitations of this success as scientists grow increasingly familiar with these treatments, and that, with the increasing duration of patients on treatment regimens, new complications arise. Overall, the rise of these topics indicates a significant change in HIV/AIDS research as it matured beyond an initial interest in identification to one of containment and treatment.

Topic networks

We now turn to how the field's topic structure further provides evidence and implications of HIV/AIDS research's maturation from an early focus on identification to one more concerned with the cost-effectiveness research, including prevention and treatment. The following network maps provide strong evidence of how maturation has influenced the extent of integration in HIV/AIDS research. We pay particular attention to how the product of this particular research program, or the content of research on HIV/AIDS, shows signs of disciplinary integration (and of what form), or whether HIV/AIDS research appears more bounded by disciplines. In Figs. 3-6, we present snapshots for a sequence of 5 year periods representing the topic overlap networks as described above.

Figure 3 presents the topic network slice for the first time period, 1990–1994. At this time HIV/AIDS research is a still-as-yet emerging field. There is a lot of strong overlap

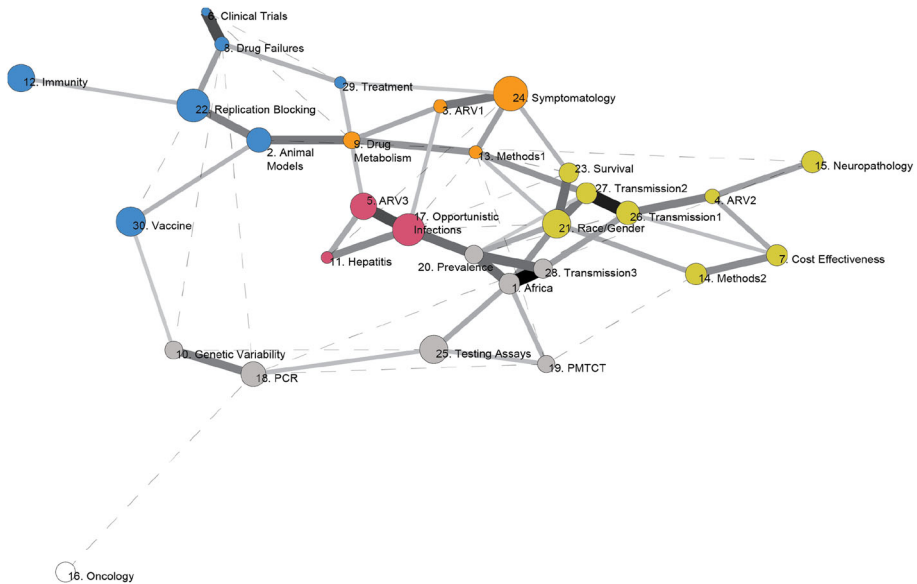


Fig. 3 1990–1994 topic network. *Note* The solid black lines represent the network for the current time period. The width of the lines corresponds to weight of overlap. To offer a picture of the dynamic process of knowledge production, the long-dashed lines indicate connections that appear in the next time period, and dotted lines represent connections that existed in the prior time period, but are no longer present. Node color is based on Louvain communities and node size is determined by the proportion of the topic within this time period. (Color figure online)

amongst the topics and a weak “division of labor” at this stage as indicated by the content of the research. No dramatic fissures exist within the network providing strong evidence of interdisciplinarity. However, even during this slice, there appear to be three rough subgroups. These subgroups do not map neatly onto the explicit divides within the HIV/AIDS research field (basic research, clinical research, and behavioral science; see Adams and Light 2014), nor do they map onto the divides that emerge in later time slices. Rather, we see three loosely bounded groups of topics pertaining to substantive issues: researchers studying the disease itself (northwestern portion of the graph, colored blue), and researchers studying people’s encounters with the disease (central/southern portion of the graph). Prevalence is central among this second cluster, and was the open question at the time as scientists tried to determine how widespread the disease was and would become. The latter group is further split between those topics concerned with HIV prevention and spread (eastern portion of the cluster, yellow cluster) and those concerned with people living with HIV/AIDS (PLQWHA, e.g., co-infection; western portion of the cluster, orange/red clusters). Interestingly, preventing mother-to-child transmission of HIV (PMTCT, gray cluster) is separate from other forms of transmission (all located in the yellow cluster) at this stage, consistent with the rather undeveloped understanding of PMTCT during this period. While researchers in this relatively nascent field produced research in a somewhat organized fashion, indicated by these loose subgroups, these subgroups appear to be based on specific research questions and not bounded by disciplines. Moreover, the topic network is very connected, with the exception of one isolate: oncology (topic 16), providing further evidence of interdisciplinarity. This slice is

consistent with a new problem-based research program, which is more likely to begin with interdisciplinarity as early-adopters attempt to make sense of little-known areas. The case of HIV/AIDS research presented here suggests that disciplinary boundaries may be easier to cross in these earlier moments.

In the next network slice (Fig. 4, 1995–1999), the integration characteristic of the early period starts to dissipate and fissures begin to appear. These fissures only grow more pronounced in later slices. Specifically, the primary split between bench (southwestern sector of the graph, blue/gray clusters) and behavioral (eastern sector of the graph) sciences arises. As clusters become more pronounced, the topic structure may become more likely to align with classic substantive divides (e.g. bench, clinical, sociobehavioral) within research on HIV/AIDS. These substantive divides represent clusters of similar disciplines and approaches to understanding HIV/AIDS and other sociomedical problems. While prior analyses have used these substantive divides more formally (see Adams and Light 2014), we use these divides to cross-check changes in the topic structure over time. In this slice, a very clear PLWHA cluster (northern sector, orange cluster) evolves consistent with the maturation of the epidemic: People are starting to live a little longer with the disease due to treatment and more people are known to be infected. Therefore, the literature is more focused on these concerns, such as opportunistic infections (topic).

At this stage, researchers throw the weight of scientific strategy behind understanding the details of the disease (clustered in the southwestern sector, blue cluster). Highly connected topics detection and treatment form a new core among bench sciences—centered around topics of treatment, clinical trials, drug failures and detection (PCR, testing assays, genetic variability, gray cluster), with weaker connections to other bench science topics pertaining to the structure and behavior of the virus itself—(red cluster) including such topics as vaccines, animal modeling and replication blocking.

In addition to the increasingly concentrated focus on the disease itself, scholars similarly consolidate focus around two sets of behavioral topics—which have strong conceptual overlaps in the etiology and development of HIV/AIDS which is also reflected in the

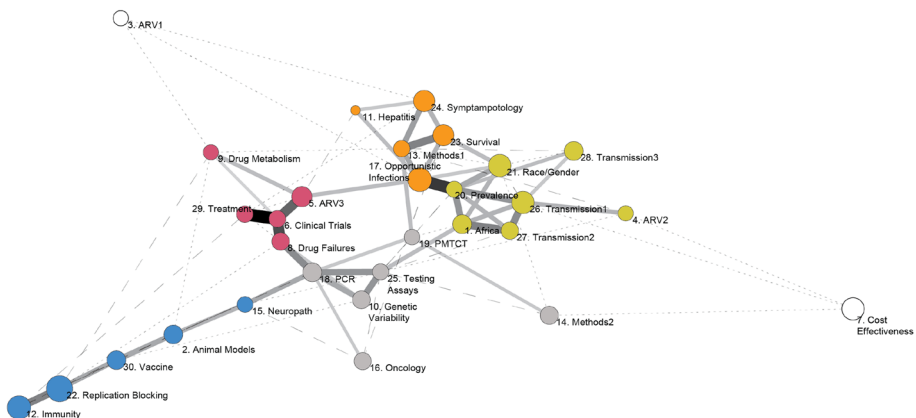


Fig. 4 1995–1999 topic network. *Note* The solid black lines represent the network for the current time period. The width of the lines corresponds to weight of overlap. To offer a picture of the dynamic process of knowledge production, the long-dashed lines indicate connections that appear in the next time period, and dotted lines represent connections that existed in the prior time period, but are no longer present. Node color is based on Louvain communities and node size is determined by the proportion of the topic within this time period. (Color figure online)

overlapping structure between these sets of topics: note the presence of a north/south divide between clusters on the eastern side of the graph (colored yellow and orange respectively), but also that there are several strong links across those sectors. These behavioral topic clusters roughly split along lines of PLWHA (identifying and treating their disease career, orange) to the north, and prevention/surveillance to the south (yellow). While there is a clear split between these clusters, it is two topics at the core of each that share the strongest overlap in this period—the opportunistic infections topic in the former and prevalence topic in the latter. This also reflects the increasing recognition of “treatment as prevention” that took note of the importance of taking care of those already infected as a means of understanding the routes of and preventing further infections from occurring. Moreover, the transmission and surveillance topics (yellow cluster) began devoting additional attention to different patterns of infection; e.g., with racial disparities more closely associated with the IDU and sexual transmission topics, while the Africa topic is also closely related to the sexual transmission topic, but also the “high risk” transmission topic, reflecting this period’s focus on accounting for the disproportionate prevalence in sub Saharan Africa—at least in part through understanding the roles of commercial sex work and higher prevalence of other (untreated) sexually transmitted infections.

In sum, we see more specialized questions begin to occur during this time period. The content of HIV/AIDS research remains relatively well-bounded and appears to be quite interdisciplinary, but we see early signs of a shift away from interdisciplinarity as measured by the content of the research itself. The network has grown visibly less connected and more organized around traditional disciplinary divides. This retrenchment only grows as HIV/AIDS research further matures in subsequent periods.

Figure 5 presents the topic network slice for 2000–2004. During this period “answered” questions emerge. Researchers by this time have uncovered the difference between HIV and other retroviruses (Harden 2012: 55–68). Researchers also have developed an understanding of opportunistic infections. At the same time, new anti-retroviral therapy (ART) regimens are being developed along with a greater understanding of PLWHA. A testing strategy is largely in place at this point. With the exclusion of these topics, the bench sciences appear to be more consolidated in this network (see the southeastern, blue, cluster) than in previous ones. Perhaps the most striking difference between this period (and the next) compared to the first is that now virtually all of the strongest observable ties are within clusters rather than across them. While the pattern between clusters in each period is marked by a few ties between seemingly loosely related topics (e.g., opportunistic infections and survival here), those cross-cluster links are much more often the weaker of the observed links, whereas in Fig. 3, the cross-cluster links were often among the strongest observed in the period (see opportunistic infections and prevalence, or animal models and drug metabolism). At this stage, the PMTCT topic bridges the bench (southeast cluster, blue) and behavioral topics (north cluster, yellow/orange), which is consistent with the growing interest in and understanding of mother-to-child transmission. Specifically, researchers grew confident that they could develop pharmacological strategies to prevent this form of transmission. For example, researchers developed an understanding of the role that viral-lode played in both mother-to-child transmission and other transmission routes. This reflected extensive efforts focused on solving a problem that was thought to be both “solvable” and disproportionately affecting “blameless” victims, the latter being a persistent undercurrent in the literature on HIV/AIDS.

The final network slice, 2005–2010 (Fig. 6), indicates the further separation between bench, clinical, and behavioral sciences. Little evidence of the initial integration exists by

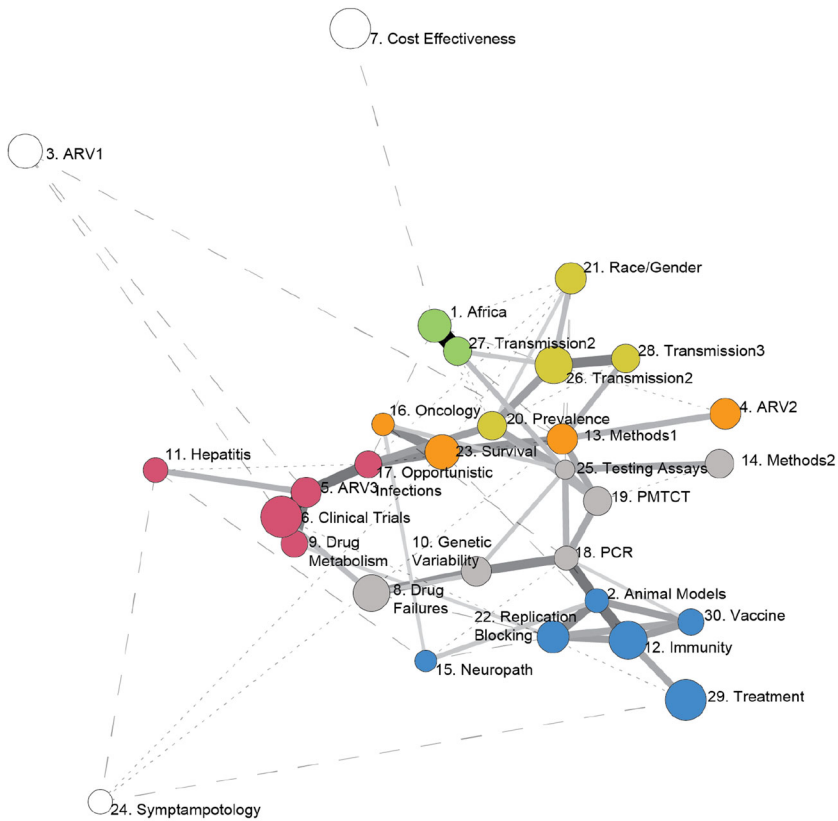


Fig. 5 2000–2004 topic network. *Note* The solid black lines represent the network for the current time period. The width of the lines corresponds to weight of overlap. To offer a picture of the dynamic process of knowledge production, the long-dashed lines indicate connections that appear in the next time period, and dotted lines represent connections that existed in the prior time period, but are no longer present. Node color is based on Louvain communities and node size is determined by the proportion of the topic within this time period. (Color figure online)

this time period. The PMTCT topic has moved even closer to the bench side (southwest, blue/green/orange clusters). This shift results from the development of nevirapine and related drugs and the “solving” of the mother-to-child transmission problem (Hankins 2000; Dabis and Ekpin 2002). This development does not indicate that mother-to-child transmission no longer occurred, but that prevention of this type of transmission is increasingly well-understood, and was often considered a clinical solution, not something left for the mother to pursue (i.e., is not in the domain of behavioral scientists) (Angotti et al. 2011). At this stage, PMTCT sits between the new divide between the bench and clinical sciences (northwest cluster), and is completely disconnected from the behavioral cluster (northeast cluster). By this time an emerging coherence of ARV-related topics appears (northwest cluster, red), which stands independent from vaccine or virus related topics (southwest cluster, orange/blue/green clusters). This reflects the positioning of ARV development as a purely bio-medical topic area, providing one specific area that has grown less integrated over time. Interestingly, the US-focused demography topic (Race/Gender,

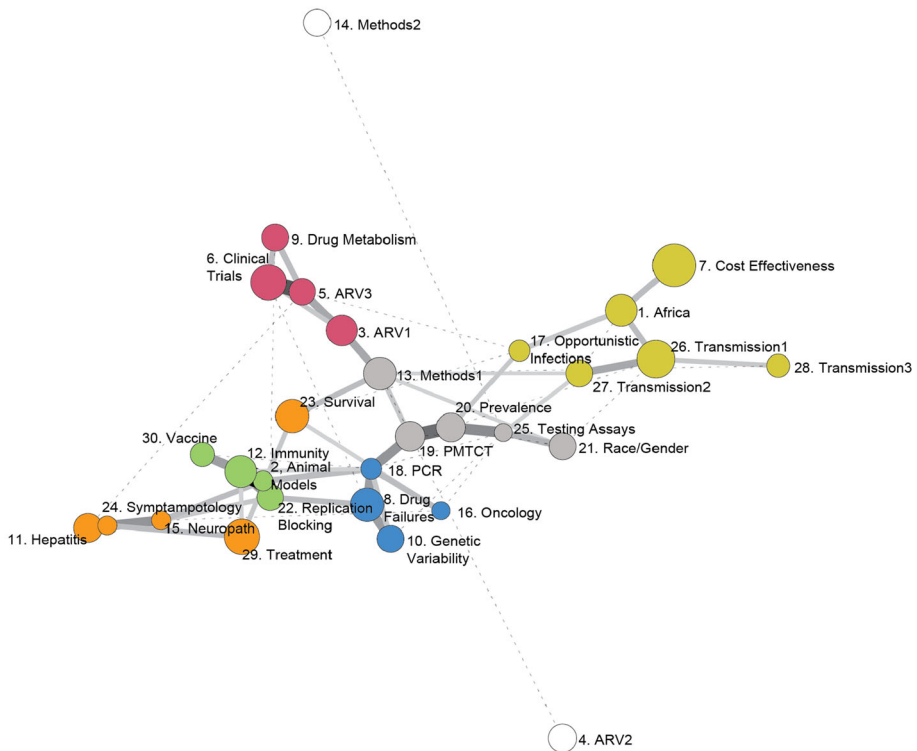


Fig. 6 2005–2010 topic network. *Note* The solid black lines represent the network for the current time period. The width of the lines corresponds to weight of overlap. To offer a picture of the dynamic process of knowledge production, the long-dashed lines indicate connections that appear in the next time period, and dotted lines represent connections that existed in the prior time period, but are no longer present. Node color is based on Louvain communities and node size is determined by the proportion of the topic within this time period. (Color figure online)

part of the central gray cluster) has increased in centrality during this final slice consistent with a recent influx of attention to HIV disparities within the United States with specific attention to particular populations showing increases of new infections e.g., African American women, and the Washington DC area.

A distinct split occurs between the bench and the behavioral science by the 2005–2010 time period. Similar to the previous period, but even more pronounced, the network is now sparser and fewer connections span the more disciplinary research approaches. As HIV/AIDS research has matured, the organization of knowledge production shifted from interdisciplinary to multidisciplinary perspective. Fewer research papers reach across firm disciplinary boundaries. This model of dynamic knowledge production points to structural holes in the topic network that may be beneficial for future HIV/AIDS researchers. For example, there may be good reason to re-engage interdisciplinary questions that connect the behavioral and bench sciences especially as advances in vaccine development bear fruit. If bench scientists actually succeed in developing a vaccine, behavioral scientists may be influential in developing strategies for distributing it. While early stages of research may require pronounced interdisciplinarity (akin to what was observed in Fig. 3) when, for

example, epidemiological questions about identifying at-risk populations intersect with clinical concerns about disease identification. However, more advanced problem-areas may require more multidisciplinary approaches as specialized questions rise to the fore. At the same time, and perhaps less optimistically, more mature problems may also have more competition over greater resources, which likely also solidifies traditional boundaries. In either case, we gain a better understanding of potential areas of promise and concern when looking at research fields from a dynamic perspective, accounting for *why* we see the pattern we do in this latter period, in part, by examining the prior organizational patterns it emerged from.

In sum, a dynamic view of the interdisciplinary work central to HIV/AIDS research tells a more complicated story than one of simple integration. In early stages of development, HIV/AIDS research appears topically integrated with high overlap between seemingly disparate content areas. This characteristic conforms with the image of researchers struggling to come to terms with a new problem and certain openness towards divergent perspectives. Over time this high level of integration weakens and traditional divides between the clinical, bench, and behavioral sciences appear as large questions are resolved and more detailed and specialized questions take center stage.

Conclusion

Tracing the history of HIV/AIDS research through its content allows for the evaluation of the dialogue between research scientists seeking to understand and mitigate its spread. It has long been understood that solutions to the HIV/AIDS epidemic require both social and biomedical research. Our results indicate that the field has changed in significant ways exhibiting signs of maturation moving from disease identification to more specialized concerns such as cost effectiveness. At the same time, HIV/AIDS research appears to have grown more multidisciplinary also consistent with increased specialization, after an initial period of more interdisciplinary integration. In other words, after an initial period of high cross-fertilization and joint problem-solving, the researchers returned to their silos working on overlapping topics in a more disciplinary fashion.

As evidenced by the case of HIV/AIDS research, integrated programs are dynamic and situate themselves into different states of organization over time. Challenging an atomistic conception of interdisciplinarity, our approach encourages a dynamic understanding of integrative research projects. As this model is carried forward, an important question will be what sorts of projects demonstrate evidence of similar (or different) trajectories. Are the types of projects that evolve into transdisciplinary (“boundary-less”) phenomenon likely to emerge from differing formative states? It also provides a means for examining more traditional questions, like how the boundaries within (sub-)disciplines evolve over time (Leahey and Moody 2014). For example, sociology’s historical evolution involves numerous points of strong connection to other disciplines, such as criminology and social work, while at the same time trying to remain a loosely bounded discipline (Moody and Light 2006).

The HIV/AIDS case presented here provides a template for examining questions of many-disciplined science as a dynamic model of integration. As such, our approach moves away from compartmentalized conceptualizations of non-disciplinary states, and avoids assuming overly deterministic linearity in processes of disciplinary integration. Moreover, it highlights that solutions to social problems are unlikely to arise from either disciplinary

or interdisciplinary approaches along, but rather from unique trajectories that combine specialization and cross-pollination.

Understanding how science evolves is essential for evaluating its successes and failures. In this way, the sociology of science can have a direct effect on the amelioration of both social and medical problems by providing methods for analyzing science and policy. The computational approach advanced in this study is one way to construct a temporal picture of scientific research. Computational approaches are helpful because they facilitate the incorporation of large amounts of data and allow us to “read”—in a very particular way—the work of thousands of scientists. These “satellite” images help us to understand a number of scientific processes, such as when scientific consensus has developed around a particular issue or the state of interdisciplinarity in a particular research field. A holistic image of science requires a multidimensional approach that considers both broad pictures and more local processes. Conceptualizing interdisciplinarity as dynamic and not an idealized end state is an important step towards building a critical understanding of how modern approaches to social and medical problems are organized.

References

- adams, j., & Light, R. (2014). Mapping interdisciplinary fields: Efficiencies, gaps & redundancies in HIV/AIDS research. *PLoS ONE*, 9(12), e115092.
- Angotti, N., Dionne, K. Y., & Gaydos, L. (2011). An offer you can't refuse? Provider-initiated HIV testing in antenatal clinics in rural Malawi. *Health Policy and Planning*, 26, 307–315.
- Bayat, A. (2002). Science, medicine, and the future: Bioinformatics. *British Medical Journal*, 324, 1018.
- Bettencourt, Luis M. A., Kaiser, D. L., Kaur, J., Castillo-Chavez, C., & Wojick, D. E. (2008). Population modeling of the emergence and development of scientific fields. *Scientometrics*, 75(3), 495–518.
- Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55(4), 77–84.
- Blei, D. M., & McAuliffe, J. D. (2010). Supervised topic models. <http://arxiv.org/pdf/1003.0783v1.pdf>.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *The Journal of Machine Learning Research*, 3, 993–1022.
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), P10008.
- Boyack, K. W., Klavans, R., & Börner, K. (2005). Mapping the backbone of science. *Scientometrics*, 64(3), 351–374. doi:10.1007/s11192-005-0255-6.
- Burt, R. S. (2004). Structural holes and good ideas. *American Journal Sociology*, 110, 349–400.
- Chang, J. (2012). *Ida: Collapsed Gibbs Sampling Methods for Topic Models*. R: CRAN. <http://www.R-project.org/>.
- Chang, J., Gerrish, S., Wang, C., Boyd-Graber, J. L., Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. *Advances in neural information processing systems*, 22, 288–296.
- Choi, B. C., & Pak, A. W. (2006). Multidisciplinarity, interdisciplinarity and transdisciplinarity in health research, services, education and policy: 1. Definitions, objectives, and evidence of effectiveness. *Clinical and Investigative Medicine. Medecine Clinique et Experimentale*, 29, 351–364.
- Couch, S. R. (2004). A tale of three discourses: Doing action research in a research methods class. *Social Problems*, 51, 146–153.
- Crane, D. (1972). *Invisible colleges: Diffusion of knowledge in scientific communities*. Chicago: University of Chicago Press.
- Dabis, F., & Ekpin, E. R. (2002). HIV-1/AIDS and maternal and child health in Africa. *The Lancet*, 359(9323), 2097–2104.
- DiMaggio, P., Nag, M., & Blei, D. M. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of US Government arts funding. *Poetics*, 41(6), 570–606.
- Dubrow, J. K. (2011). Sociology and American studies: A CSE study in the limites of interdisciplinarity. *The American Sociologist*, 42, 303–315.
- Epstein, S. (1995). The construction of lay expertise: AIDS activism and the forging of credibility in the reform of clinical trials. *Science, Technology and Human Values*, 20, 408–437.

- Epstein, S. (1996). *Impure science: AIDS, activism, and the politics of knowledge*. Berkeley: Univ of California Press.
- Frickel, S., Albert, M., & Prainsack, B. (Eds.) (2016). *Investigating interdisciplinary research: Theory and practice across the disciplines*. Rutgers University Press.
- Fujimura, J. H., & Chou, D. Y. (1994). Dissent in science: Styles of scientific practice and the controversy over the cause of AIDS. *Social Science and Medicine*, 38(8), 1017–1036.
- Gieryn, T. F. (1983). Boundary-work and the demarcation of science from non-science: Strains and interests in professional ideologies of scientists. *American Sociological Review*, 48, 781–795.
- Gieryn, T. F. (1999). *Cultural boundaries of science: Credibility on the line*. Chicago: Chicago University Press.
- Gondal, N. (2011). The local and global structure of knowledge production in an emergent research field: An exponential random graph analysis. *Social Networks*, 33(1), 20–30.
- Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Science*, 101(S1), 5228–5235.
- Grun, B., & Hornik, K. (2011). Topicmodels: An R package for fitting topic models. *Journal of Statistical Software*, 40(13), 1–30.
- Hackett, E. J., & Rhoten, D. R. (2009). The Snowbird Charrette: Integrative interdisciplinary collaboration in environmental research design. *Minerva*, 47(4), 407–440. doi:10.1007/s11024-009-9136-0.
- Hadorn, G.H., Biber-Klemm, S., Grossenbacher-Mansuy, W., Hoffmann-Riem, H., Joye, D., Pohl, C., Wiesmann, U., & Zemp, E. (Eds.). (2008). The emergence of transdisciplinarity as a form of research. In *Handbook of Transdisciplinary Research* (pp. 19–39). Netherlands: Springer.
- Hankins, C. (2000). Preventing mother-to-child transmission of HIV in developing countries: Recent developments and ethical implications. *Reproductive Health Matters*, 8(15), 87–92.
- Harden, V. A. (2012). *Aids at 30: A history*. Washington, DC: Potomac Books.
- ISI, Thompson Reuters. (2012). *Journal Citation Reports®*, Science edition.
- Jacobs, J. A. (2013). *In defense of disciplines: Interdisciplinarity and specialization in the research university*. Chicago: University of Chicago Press.
- Jacobs, J., & Frickel, S. (2009). Interdisciplinarity: A critical assessment. *Annual Review of Sociology*, 35, 43–65.
- Klein, J. T. (2001). “The discourse of transdisciplinarity: An expanding global field. In J. T. Klein, W. Grossenbacher-Mansuy, R. Haberli, A. Bill, R. W. Scholz, & M. Welti (Eds.), *Transdisciplinarity: Joint problem solving among science, technology and society*. Basel: Birkhauser.
- Klein, J. T. (2010). A taxonomy of interdisciplinarity. In R. Frodeman, J. T. Klein, & C. Mitcham (Eds.), *The Oxford handbook of interdisciplinarity*. New York, NY: Oxford University Press.
- Knorr-Cetina, K. (1999). *Epistemic cultures: How the sciences make knowledge*. Cambridge, MA: Harvard University Press.
- Leahey, E., & Moody, J. (2014). Sociological innovation through subfield integration. *Social Currents*, 1, 228–256.
- Leydesdorff, L., & Hellsten, I. (2006). Measuring the meaning of words in contexts: An automated analysis of controversies about ‘Monarch butterflies’, ‘Frankenfoods’, and ‘stem cells’. *Scientometrics*, 67(2), 231–258.
- Leydesdorff, L., & Schank, T. (2008). Dynamic animations of journal maps: Indicators of structural changes and interdisciplinary developments. *Journal of the American Society for Information Science and Technology*, 59(11), 1810–1818.
- Light, R. (2014). From words to networks and back: Digital tex, computational social science and the case of presidential inaugural addresses. *Social Currents*, 1(2), 111–129.
- Lucio-Arias, D., & Leydesdorff, L. (2009). The dynamics of exchanges and references among scientific texts, and the autoepoiesis of discursive knowledge. *Journal of Informetrics*, 3(3), 261–271.
- Marshall, E. A. (2013). Defining population problems: Using topic models for cross-national comparison of disciplinary development. *Poetics*, 41(6), 701–724.
- McFarland, D. A., Ramage, D., Chuang, J., Heer, J., Manning, C. D., & Jurafsky, D. (2013). Differentiating language usage through topic models. *Poetics*, 41(6), 607–625.
- McKenna, S. L., Muyinda, G. K., Roth, D., et al. (1997). Rapid HIV testing and counseling for voluntary testing centers in Africa. *AIDS*, 11(Suppl 1), S103–S110.
- Mohr, J. W., & Bogdanov, P. (2013). Introduction—topic models: What they are and why they matter. *Poetics*, 41(6), 545–569. doi:10.1016/j.poetic.2013.10.001.
- Moody, J., & Light, R. (2006). A view from above: The evolving sociological landscape. *The American Sociologist*, 37, 67–86.

- NAS, National Academy of Sciences, National Academy of Engineering, & Institute of Medicine and Committee on Facilitating Interdisciplinary Research. (2005). *Facilitating interdisciplinary research*. Washington, DC: National Academies Press.
- Newman, E. A., Guest, A. B., Helvie, M. A., Roubidoux, M. A., Chang, A. E., Kleer, C. G., & Diehl, K. M. (2006). Changes in surgical management resulting from case review at a breast cancer multidisciplinary tumor board. *Cancer*, 107, 2346–2351.
- Nicolescu, B. (2002). *Manifesto of transdisciplinarity*. Albany, NY: State University of New York Press.
- Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, 14(3), 130–137.
- Porter A. L., Cohen A. S., Roessner J. D., Perreault M. (2007). Measureing researcher interdisciplinarity. *Scientometrics*, 72, 117–147.
- Powell, W. W., White, D. R., Koput, K. W., & Owen-Smith, J. (2005). Network dynamics and field evolution: The growth of interorganizational collaboration in the life sciences. *American Journal of Sociology*, 110(4), 1132–1205.
- Ramage, D., Rosen, E., Chuang, J., Manning, C. D., McFarland, D.A. (2009). Topic modeling for the social sciences. in *Workshop on Applications for Topic Models, NIPS*.
- Salter, L., & Hearn, Alison M. V. (1997). *Outside the lines: Issues in interdisciplinary research*. Kingston, Ontario: Queens Univ School of Policy.
- Small, H. (2010). Maps of science as interdisciplinary discourse: Co-citation contexts and the role of analogy. *Scientometrics*, 83(3), 835–849.
- Steyvers, M., & Griffiths, T. (2007). Probabilistic topic models. *Handbook of Latent Semantic Analysis*, 427(7), 424–440.
- Westin, T., & Stalfors, J. (2008). Tumour boards/multidisciplinary head and neck cancer meetings: Are they of value to patients, treating staff or a political additional drain on healthcare resources? *Current opinion in otolaryngology & head and neck surgery*, 16, 103–107.