

# the **phy**lophilosophy project

an introduction to the future of research in philosophy

- 1 // INTRODUCTION
- 2 // THE NEED FOR PHYLO
- 5 // HOW PHYLO WORKS
- 8 // HOW YOU CAN USE PHYLO

# introduction

**WE'RE ABOUT TO LAUNCH  
A NEW TOOL** at The  
Graduate Center of the City  
University of New York that will  
change the face of research

in philosophy. We've dubbed it "The Phylosophy Project" ("Phylo" for short) because, like taxonomy, it illustrates the origins of contemporary philosophy by looking at historical relationships between individuals, institutions, and ideas.

This document will introduce Phylo in several steps:

- // First, we'll talk about existing research tools and the need for Phylo.
- // Second, we'll describe how Phylo works and the data that drives it.
- // And finally, we'll talk about how you can use Phylo in your teaching and research and where Phylo can go in the future.

"The history of philosophy is to a considerable extent the history of groups. Nothing abstract is meant here—nothing but groups of friends, discussion partners, close-knit circles that often have the characteristics of social movements." *Randall Collins, The Sociology of Philosophies*

# the need for phylo

## RESEARCH IN PHILOSOPHY

**R** has changed dramatically with the rise of the internet. Sources that were once available only in hard copy bibliographies and citation lists are now cataloged in expansive databases—searchable, savable, and, in an age of Google Books and JSTOR, offering access to primary documents in a matter of minutes. But for all their impressive uses of digital technologies, current research tools are quite limited in the search capabilities they offer and types of information they contain: keyword queries of major publications, such as books and journal articles. This is not the only way of doing research, or maybe even the most intuitive.

It's because of these limitations that we're creating Phylo, a tool that illustrates connections between individuals, institutions, and ideas. Before talking about the new research possibilities Phylo presents, let's examine current research tools and the limits they impose on our ability to discover sources.

## EXISTING TOOLS AND THEIR LIMITS

Current research tools are concerned with cataloging a few kinds of publications. They relate those publications to each other using keywords, or descriptors and personal names that are attached to records and documents. To find sources, users match their areas of interest to these pre-defined keywords. The better the keywords and the better the match, the better the results returned. But plenty of things can go wrong along the way. A keyword that's too broad will return hundreds of results, a narrow one will leave out relevant sources, and sometimes, keywords are just plain unintuitive. The only way to find this out is to start searching, hope that you've used the right keywords, see what they return, determine if it's good enough, and

**“Computers are now critical to the advancement and scrutiny of many philosophical ideas. We live in new times—a computational age....Philosophy has taken a computational turn.”** *APA Newsletter on Philosophy and Computers, Fall 1997*

continue searching until it is. This method has obvious drawbacks in efficiency and accuracy, especially when databases don't publish information about the completeness of their records—and most don't.

But there's a much bigger problem with this process, and it concerns the way we're encouraged to research. Current research tools are built to deliver publication citations. But no matter how many citations they return for any search or how good those citations are, these tools tell you next to nothing about the *people* whose publications you're searching—for instance, what their connections are to other people or institutions. You have no idea whether this person got a PhD from Fly By Night Academy or Old Prestigious U, much less whether they got it in the days of Quine or the days of Rawls. No matter how many sources a database contains or how good its keywords are, users still

need to input the *right* keywords, and they still need to sift through and supplement their results with—you guessed it—*more* citations.

All of this requires some knowledge of *context*: how an idea arose in the literature, how it's connected to other ideas, which other sources might prove relevant for research. Of course, if you're an expert on, say, direct reference theory, that's no problem; you know that information anyway. You know when the theory first recognizably emerged, you know the exchanges that happened among leading figures, you know the passages by earlier authors that prefigured the theory, you know which variants of the theory were developed in response to objections, and you even know the literature on *these* subsequent theories and the history of *their* development. But in an age of increasing specialization, few people can claim to have this kind

perspective. Assembling a more complete picture requires weeks, even years of intensive research. It takes careful study of significant trends, historical influences, and, of course, personal connections between major (and minor) figures. None of this appears on the surface of current research tools, with their focus on keywords and sources, rather than context and chronology.

Imagine if some researcher tried to piece together *your* career based on the few citations contained in current tools. Would these sources say anything about who your influences have been? How and where they've been clustered together over time and place? How your work (or someone else's) on one topic has shaped your views on others? You know how important this contextual data is to understanding your own career, yet there's almost no access to the same information about others who

have worked and taught before you and shaped your philosophical development. This is a serious limit on research that's imposed, not by the information itself, but by the methods and content of current research tools.

More important than what content and capabilities current tools leave out is which potential users they exclude. Any researcher who's not already an expert—or willing to become an expert—on some topic can't easily access large portions of knowledge. This includes

most undergraduates, as well as professors from other disciplines, not to mention professionals, the media, and the general public. Perhaps it's no surprise, then, that contemporary philosophy is often seen as isolated and that the recent work of the American Philosophical Association on Public Philosophy has been praised as pioneering and important for the field.

At a more basic level, the current model leaves out anyone who doesn't have access to leading databases, most of which carry annual subscription fees. This includes researchers unaffiliated with institutions, as well

**PHILOSOPHY IS, FIRST AND FOREMOST, A CLASH OF IDEAS. BUT BEHIND THOSE IDEAS ARE PEOPLE—people who share ideas, give feedback, talk to each other, people who teach and research together, who advise others, people who write books and articles that move ideas forward. Until now, our research tools have been limited to abstract topics and simple keywords. But what if you could see the people behind an idea? See how they were connected through place and time? What if you could watch an idea develop over time? Or track shifts in the field across several decades? We think you should be able to—and we think it should be free.**

of knowledge about any more than a handful of topics. Researchers need a tool that mimics contextual knowledge and makes it available for everyone.

### WHAT (AND WHO) IS LEFT OUT

For the most part, current research tools are limited to citations of books and journal articles. These are certainly the lifeblood of the field, but they may not be enough to form a complete picture of what's happening in philosophy at a particular time. After all, any given book or article is a product of its own time and hardly has the benefit of

as students and professors at institutions whose libraries can't afford the subscription fees—especially smaller institutions and community colleges. These groups need research tools as much, arguably more, than anyone.

## HOW WE WISH WE COULD RESEARCH

Research tools are evolving entities, with one modification building on another. But each stage is usually constrained by past ones. Current tools began as simple electronic databases of records. Then users were given the ability to search by more complex keyword strings, then access to the records themselves. Now, social bookmarking tools like del.icio.us and Digg allow users to tag information themselves and share those tags with others.

But imagine you could design a research tool from the ground up. Imagine you weren't limited to just a few sources of information and a single way of accessing them. You'd probably start by including the same types of data about others that you know about yourself: where you've studied; who you've been connected to; what topics you've worked on over time; how these people, places, and publications have influenced each other across your career.

This contextual information *is* available, but it's scattered across library archives, administrative reports, and even publications. Just think, for example, of everything that a list of dissertation committee members tells you about why an author took *that* particular approach to some problem or why that author included a chapter on one topic, but left out a discussion of another. At present, there's no single tool that collects this kind of contextual information (there's not even a collection of them that accomplishes this), much less ties it back to the primary sources that fuel our research. This lack underscores the need for Phylo, a new online tool that examines historical connections in philosophy—free and accessible to all researchers. //

**THE MISSING LINK IN RESEARCH** You'd probably agree that the people around you—your colleagues, your students, your coauthors and editors—have had a huge influence on your work. You'd probably also agree that these are the *last* things you search for when you research another philosopher or even an idea. Why? Chances are it's because there's just no good tool that captures this information—until now. With Phylo, you'll be able to see these relationships in a simple, intuitive way. Then you can fill out your research with new context, or track down new primary sources, or show your students how ideas have emerged over time, or...

# how phylo works

**A** TOOL LIKE PHYLO that gives context to current understandings and creates new paths for research could be invaluable. But the amount of data involved here is huge. There's a wealth of professional data to capture—faculty appointments, attendance at institutions, dissertation supervision and dissertation committee work—not to mention additional publication data from books and journals. It all needs to be searchable and it all needs to lead back to primary sources. Once you realize how many people and documents we're talking about and how many centuries they span, you might start to think there's *too* much data here to assimilate.

We'd tend to agree, except for the fact that all of us make sense of information like this every day. We routinely keep track of people we know, where they are, what they're doing, and how they're connected. Understanding social networks is second- (maybe even first-) nature to us. That's why we've built Phylo's data displays around the familiar concepts of individuals, institutions, and ideas. It's an intuitive way to process large amounts of data, and a natural model for expanding data for years to come.

## INDIVIDUALS, INSTITUTIONS, AND IDEAS

Phylo's main categories are drawn from people, places, and publications—all real-world objects that we have common ways of representing.

Individuals are connected to each other through different kinds of relationships. These relationships form larger webs, with each individual serving as a node connected to other individuals in different ways. It's now common to visualize these connections as social networks, but there's a much older pattern we're used to tracking: family trees. These are historical, graphical representations of how people are connected by birth or marriage. These two relationships are quite similar to advisor–advisee pairings and professional peers. This similarity gives us a very basic way of representing people using historical networks of connections.

Institutions are historical entities. They grow and change and shift over time, but they remain the single thread that connects all of their members and activities together. This continuity suggests a timeline with individuals and events plotted in an ordered and meaningful sequence, a visual narrative of how a department has developed over time.

Ideas are carried by books and articles, both of which can be counted—and carried to places themselves. These aspects of publications give us several ways of representing them. One way to show prominence is just to count up the number of publications associated with a particular keyword—yes, keywords are still very important—in play at any given time. If we extend this counting historically, we get a sense of how ideas grow and change. And if we add in geographical data about where people are developing those ideas and transporting them, we can plot a map that shows where ideas are being developed and transported.

Nodes, timelines, and maps aren't the only ways to represent people, places, and publications, but they are some of the most obvious, given how we represent them in our everyday lives. By combining these visualizations with the common search-and-result model, Phylo offers a contextually rich way to access primary sources. You won't need a huge background of knowledge to find the right sources; you'll be able to explore a few displays that are built around visualizations you're already comfortable with.

## THE DATA BACKBONE

Research tools are only as powerful as the data that drives them. That's why the launch version of Phylo will include dissertation and faculty information from roughly 20 North American institutions, all the way back to their first philosophy PhDs. We chose these schools based on a mix of criteria, including past and current prominence, historical longevity, number of dissertations granted, and geographical accessibility (after all, we're on a time limit and a small budget). This initial data should cover about 40% of all dissertations written in North America\* and a huge portion of very important—and currently hard-to-access—information about faculty appointments and dissertation committee service.

There are obvious ways this data could be expanded: more schools, different regions, going further back in time. We'd also like to fold in citations of books and articles, which would round out the full search capabilities of Phylo. The launch version won't contain all this, but the beauty of Phylo's electronic form is that this information can be added—quickly and by other users.

## NOW IT'S YOUR TURN: EXPANDING DATA

Phylo will give users the ability to add and edit information quickly and easily, anything from a dissertation citation to

the full range of a faculty appointment—or even a few of these at once. This model has two important advantages.

First, it distributes the database maintenance across many users, who each know a little piece of information. Most existing research tools require the hard work of a few people to update. This makes additions slow, and when funding runs out, the project is left at a standstill. Phylo, on the other hand, doesn't require a dedicated team of researchers for expansion. It only requires the knowledge of users—the *expert* knowledge, that is, of dissertation writers, colleagues, biographers, publishers, and others who have authoritative knowledge about their own projects and personal connections. By pooling these small pieces of information together, Phylo can amass a wealth of data

**THINK YOU'RE WILLIAM JAMES' GREAT GREAT GRANDCHILD? Hopefully, Phylo can tell you. We've started with a large amount of data from North American schools because they offer a standard source of information: the dissertation. But Phylo has the capacity to expand indefinitely; the only limit is gathering data. That's why user submissions are so important. With them, Phylo might even be able to tell you if you're a descendent of, say, Thomas Reid, who was once regarded as more important than Hume. We bet there are a couple of people out there who can name a few of Reid's students—and others who can name a few more. By pooling our knowledge, we can take Phylo far beyond any research tool that exists today.**

without significant staffing or financial constraints. Software engineers talk about this in terms of *scalability*—the ability to handle growing amounts of work in a graceful way—and high scalability is a crucial part of a lasting information resource.

Second, allowing users to upload their own information will keep Phylo current on new dissertations, new faculty hires, and even new publications. All the hard work goes into gathering the historical information. Once it's been collected, small updates can be made by individual users who want to make their own marks on Phylo.

\* According to estimates from Dissertation Abstracts International.

As with any user-upload resource, there needs to be a way to keep Phylo's data accurate. That's why we've created a data integrity model around the idea of peer review. Whenever a user submits data, he or she has the option of attaching documentation or providing a reference for that information. You'll be able to see that document or citation right next to the data itself. If there's any piece of data you want to challenge, you can flag it, and Phylo will send an email to the uploader asking him or her to respond with proof (if it's not already on file). In the meantime, others will see the flag and be able to decide for themselves whether or not to use the data. When something is verified, you'll see that, too, and the particular datum will

be locked down from further editing, so there's no worry about redoubling documentation efforts. In short, Phylo will display everything it contains about a piece of data so you can assess the reliability of that information yourself. All of the launch-version data will be verified, and Phylo will be able to roll back to any previous stage, if necessary.

Phylo is designed to put power in the hands of users, not only to search, but also to contribute and correct information in the database. These features make Phylo a living research tool that can expand and adapt to meet more research needs. And unlike other tools, it won't take weeks or months to release new updates. You'll see everything live, as it's added. //



# what you can do with phylo

**W**E TALKED EARLIER

about Phylo expanding opportunities to research contextual data and extending access

to nonexperts and to researchers who don't have access to databases with hefty subscription fees. Now it's time to get more specific about the ways you can use Phylo in your research and teaching. These are just a few of the things you can do with Phylo; there are probably more we haven't even thought of yet!

## DISCOVER SOURCES

Like other research tools, Phylo is designed to deliver primary source citations. We'll even try to get you as close to the texts as possible using finding aids like WorldCat.org, Google Scholar, and book retailers. But the big difference in using Phylo is that you won't just be limited to keyword searches. You'll be able to draw on people, places, and chronologies to access sources you didn't know about before or wouldn't think to include. Some of these sources may be major pieces of a debate; other, lesser-known ones may give you insight into how an idea was received, what challenges it faced, and how it was defended in the literature.

## STUDY TRENDS

There's been a lot written in the past few years about the history of twentieth-century philosophy. But so many of these stories are told through a few major figures and seminal articles. Smaller figures and minor debates are left to the side, and for good reason—until now, there's been no convenient way to search these sources.

Phylo will change the way researchers study and write about trends in the field. Say you want to research John Dewey's influence on twentieth-century ethics. (This would probably be a pretty good thing to do, given that Dewey and Tufts' *Ethics* was the leading textbook for several decades.) You could start by tracing Dewey's students at the University of Michigan, the University of Chicago,

**ARE WE LOOKING AT WEB 3.0?** We don't care much about buzzwords, but we do think Phylo is doing something really new with internet technologies. Right now, Web 2.0 applications are placing control directly in the hands of users. Instead of accessing published information, users are establishing presence and participating in content creation, making data dynamic, distributed, and rapidly updatable. If blogs, wikis, and social networking software are the nuts and bolts of our web experience, Phylo is a new power tool. Phylo takes data from a number of sources—documents, users, departments—and makes it searchable and manipulable on the fly. Users can switch back and forth between displays and data types, metadata and primary source links. This intuitive, visual approach is pushing database technology to new limits that make more data more meaningful for more users.

and Columbia University, including their dissertation topics and where they went on to teach. From there, you're on your way to dozens, even hundreds of new sources to examine and synthesize in your research.

Or here's another example. Suppose it would help to know which departments specialized in which areas throughout the twentieth century. With Phylo, you can see an entire departmental history told in terms of doctoral students and faculty appointments. Based on these figures and their topics of research, you'll get a clear and accurate picture of what's being done where and when. We'll even color code the displays according to areas of philosophy to make this information visible at a glance.

## GET YOUR STUDENTS INVOLVED

Teaching undergraduates contemporary philosophy can be challenging. For starters, there's a large number of sources, specialized terminology, and assumed background knowledge. Still, it's natural to think of twentieth-century philosophy in terms of dialogue and exchange between major figures and positions. Phylo reflects these continuities by putting ideas in the context of people, places, and other ideas. It makes this information accessible to students through displays they're already comfortable with: networks, timelines, maps, etc. Phylo can be used as a visual aid to introduce a topic, as a prompt for a research assignment, or as a general resource you can recommend to your students. Again, these are only a few possibilities, and you'll probably discover more as you begin to use Phylo yourself.

## A FRESH LOOK AT THE FIELD

By now, you should be seeing the field of philosophy in a whole new light. Soon, you'll be able to discover connections between ideas and people you didn't know before, and those connections will lead you to new people and ideas, and new connections, and so on. You'll be able to track an idea as it's developed across time, through conversation and exchange between different schools and scholars. In short, you'll be able to study familiar topics in new ways—not based on anecdotes, not based on vague impressions of how things have gone, but based on solid, empirical data from primary documents. We'll even try to provide links to those documents to further your research.

Phylo's launch-version data will cover some of the most recent work in the history of philosophy. With user submissions and further expansion, Phylo will continue to grow in size and significance. Imagine the possibilities if we could put every philosopher of the past 200, 500, even 1,000 years in the database. Imagine if we could find links from the present all the way back to the Early Moderns, or even Aquinas. Or imagine if we adapted Phylo for use in other fields; the types of data would be essentially the same, the displays would be equally meaningful. Then imagine if we created a system of keywords that linked all of this data together. The result would be a dynamic, visual catalog of human intellectual history: all of the people, all of the places, all of the ideas that have interacted with each other and moved our knowledge forward across the centuries. Phylo is a long way from doing all this, but the basic framework is there. It's just waiting for the right information from the right people. //

**the philosophy project** explores the origins of contemporary philosophy by looking at historical relationships between individuals, institutions, and ideas. These relationships are extracted from a database of primary and secondary documents and rendered using information visualization tools.

Phylo was created by David Morrow and Chris Alen Sula in the PhD/MA Program in Philosophy and the New Media Lab at The Graduate Center of the City University of New York. The project is maintained by our many users who upload data.

project website  
[www.phylophilosophy.net](http://www.phylophilosophy.net)

blog  
[www.phylophilosophy.net/blog](http://www.phylophilosophy.net/blog)

email  
[phylo@phylophilosophy.net](mailto:phylo@phylophilosophy.net)