

IDS CASE STUDY: Chapin Hall

Leveraging Chapin Hall's
Mission to Enhance Child Well-Being

by Erika M. Kitzmiller

AUGUST 2013

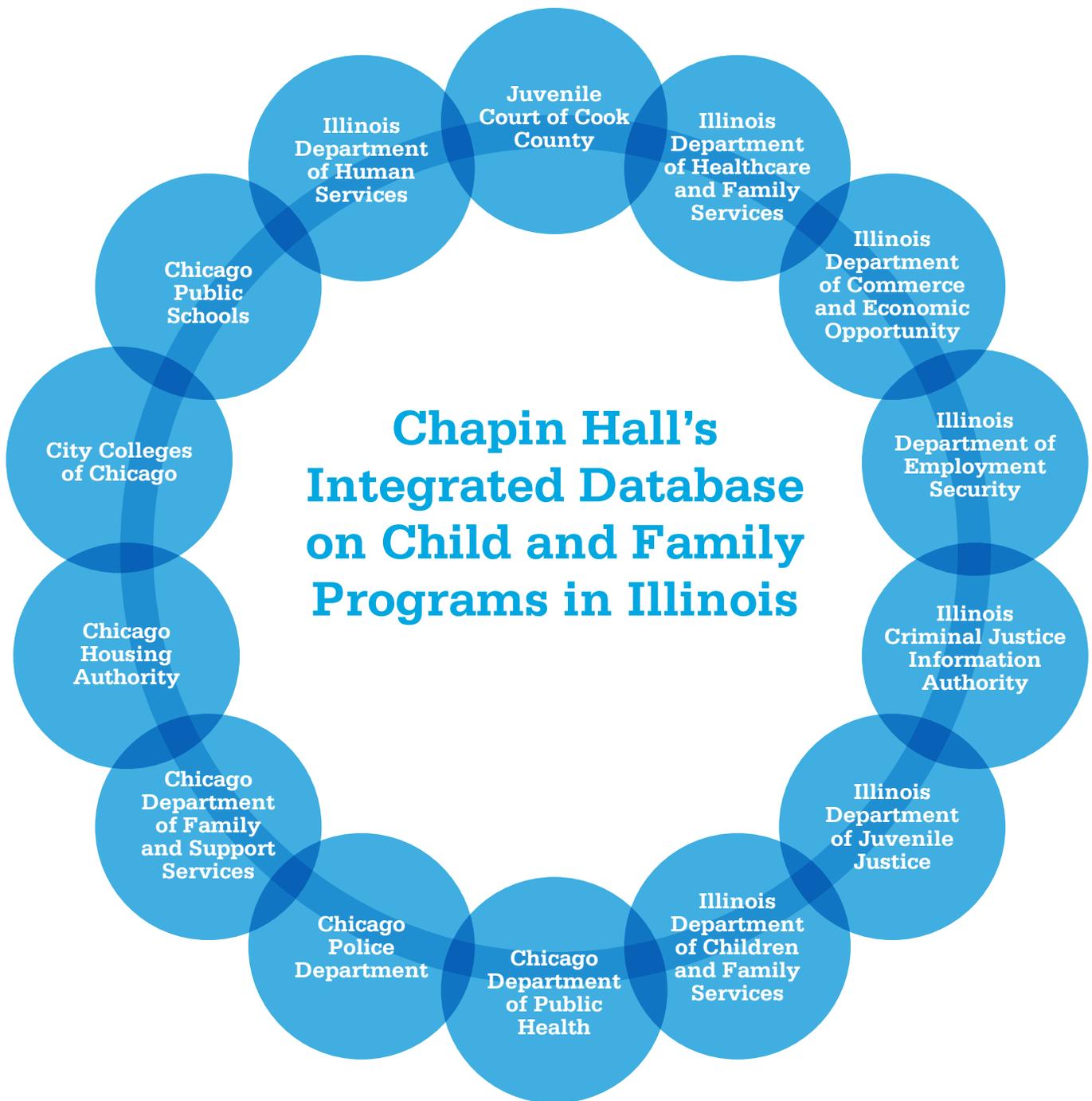


Chapin Hall’s Integrated Database on Child and Family Programs in Illinois connects administrative data on social service receipt, education, criminal and juvenile justice, employment, healthcare, and early childhood programs to provide a comprehensive picture of child and family use of publicly provided or financed service programs. Researchers who use Chapin Hall’s integrated database produce policy briefs and reports for policymakers who are interested in improving the programs that serve children and their families. Currently, the integrated database links data from a variety of agencies, such as the Chicago Public Schools, the Department of Corrections, and the Department of Family and Children Services (see figure 1 for a complete list of agencies that contribute data).

FIGURE 1. Contributing Agencies, Chapin Hall’s Integrated Database on Child and Family Programs in Illinois

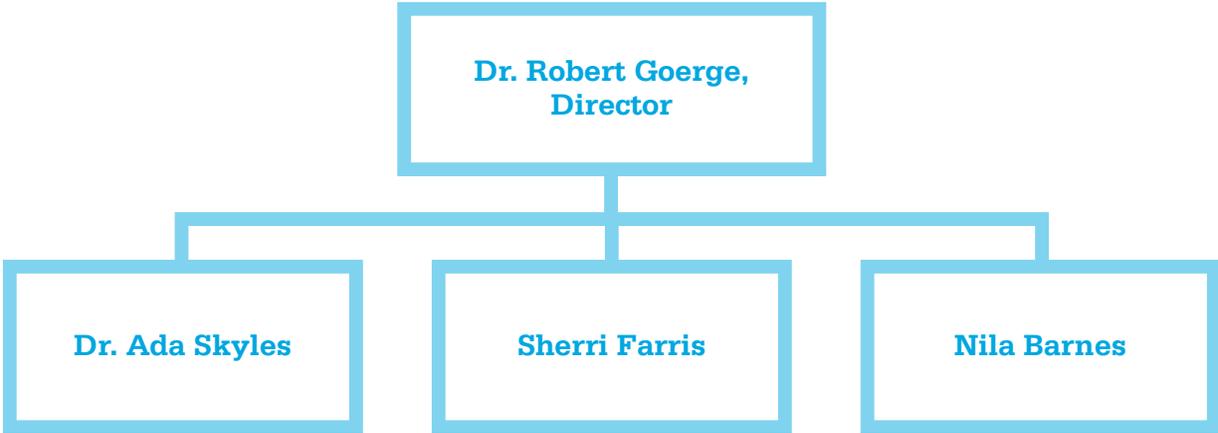
(see next page)

FIGURE 1. Contributing Agencies, Chapin Hall's Integrated Database on Child and Family Programs in Illinois



Dr. Robert M. Goerge, a Chapin Hall Senior Research Fellow, is the founder and director of Chapin Hall’s Integrated Database on Child and Family Programs in Illinois and oversees database operations and research. Nila Barnes, the director of the Illinois Data Warehouse, is the database administrator and leads all data processing. Dr. Ada Skyles, Associate Director of Chapin Hall and Research Fellow, and Sherri Farris, the data acquisition manager, are responsible for securing data sharing agreements and acquiring the data once the data sharing agreements are in place.

FIGURE 2. Chapin Hall’s Integrated Database on Child and Family Programs in Illinois Organizational Chart



During the site visit, I conducted semi-structured interviews with these four staff members as well as Dr. Lucy Bilaver, a faculty member at Northern Illinois University’s School of Nursing and Health Studies and Senior Researcher at Chapin Hall, about their roles and practices with Chapin Hall’s Integrated Database. The interviews were based on an interview protocol developed with the University of Pennsylvania’s *Actionable Intelligence for Social Policy* (AISP) team. Using standard qualitative methods, I coded these interviews for four themes—securing and maintaining legal agreements; running governance processes; data analytics management and processes; and economic and political realities to sustain operations—to understand the evolution of this integrated

data system (IDS) and the benefits that Chapin Hall’s integrated database offers to policy makers and analysts.

Like the other university IDS operations, Chapin Hall’s Integrated Database does not have a formal governance procedure in place. Instead, Goerge screens, approves, and manages any research project that utilizes the integrated database and often serves as a principal investigator on these projects. Chapin Hall staff members also enlist support from research and policy experts throughout Illinois to support database work and provide their expertise and knowledge on particular topic areas. For example, Goerge and his team works closely with policy analysts affiliated with the University of Chicago’s Consortium on Chicago School Research on educational policy issues (<http://ccsr.uchicago.edu/>). In addition, Chapin Hall staff members have collaborated with experts in the field, such as Dr. Jens Ludwig, who leads the Crime Lab at the University of Chicago and conducts random assignment studies on the prevention of criminal behavior (<http://crimelab.uchicago.edu/>).

The inclusion of agency staff as part of the research design process is one way that Goerge ensures that the integrated database research meets Chapin Hall’s mission and improves policy outcomes for children and their families.

Even though Chapin Hall does not have a formal governance body, the staff members have thought critically about the aims and goals of the integrated database. The work that uses Chapin Hall’s integrated database is both researcher-driven and agency-driven. Goerge describes how there is often a negotiation between Chapin Hall and the agency to ensure that the research has practical outcomes and improves policy and service outcomes for children and their families, while also contributing to knowledge in the field. When researchers from the University of Chicago or other affiliates approach Goerge about a particular study or idea, he or Sherri Farris, the Data Acquisitions Manager, discuss the research proposal with their contacts at the agencies that have contributed data to the database. They routinely invite agency staff to provide

their input and ideas about the research to ensure that the study meets agency needs and that its findings will address policy concerns that the agency might have. As Goerge said in his interview, Chapin Hall’s integrated database research staff members “try to mold research into something that we want to do and [that] addresses their [the agencies’] needs...we want their input so that we can figure out what might be useful to them” (Interview, Goerge, April 22).

Although there is not a formal procedure for this process, Goerge believes that this approach is directly related to Chapin Hall’s mission. Since its founding, Chapin Hall has focused on a mission of improving the well-being of children, families, and their communities. According to Goerge, research that uses the integrated database must have a direct impact on child well-being and must also be communicated to key policy makers and other stakeholders as quickly as possible. The inclusion of agency staff as part of the research design process is one way that Goerge ensures that the integrated database research meets Chapin Hall’s mission and improves policy outcomes for children and their families.

The first step in creating any IDS is developing relationships and cultivating trust with the agencies that contribute data. For the past 25 years, Goerge and his team have worked closely with state agencies and developed specific guidelines to ensure secure data use and practice and compliance with local, state, and federal mandates regarding data storage and use. Dr. Ada Skyles and Sherri Farris oversee the processes for securing and maintaining the various legal agreements that govern data use and practice for Chapin Hall’s Integrated Database on Child and Family Programs in Illinois. Skyles and Farris work closely with agency staff and lawyers to secure individual memorandum of understanding (MOU) with each agency that contributes data to the integrated database. While the MOUs have language that is specific to each agency, the agreements share several characteristics: each MOU must clearly articulate how Chapin Hall staff will retrieve the data from the agencies, how they will store the data on secure servers, and how they and others will use the data for policy-driven research. In addition, the MOUs must meet local, state, and federal regulations, such as FERPA, that govern the integrated database’s data use and practice. Skyles and

Farris believe that agency-specific MOUs are preferable to multi-agency MOUs because individual MOUs allow agency staff and lawyers to customize these agreements to meet each agency's procedures and mandates regarding data sharing, storage, and use.

Even though Skyles and Farris have developed a streamlined process for drafting and securing these MOUs, the MOU approval process often takes several months or years depending on the will of agency staff to complete them as well as the legal complexities that govern data sharing and use. To mitigate these challenges and time constraints, Skyles and Farris often draft supplementary master MOUs with the agencies. These master MOUs outline the major guidelines regarding data use and practice and ideally often span a four or five-year period. According to Skyles, these master MOUs generally outline “the process to get permission for particular projects as well as the process for bringing the data” into the database which expedites two processes: negotiating internal data analytics and moving a research proposal to an approved study (Interview, Skyles, April 22, 2013). Skyles and Farris revisit these master MOUs with agency staff on a project-by-project basis to ensure that research practice, use, and dissemination processes meet the parameters of the master MOUs. When necessary, Skyles and Farris draft project-specific MOUs to ensure compliance and clarity (Interview, Skyles and Farris, April 22, 2013). Once the MOUs are in place, Chapin Hall's integrated database staff begin a two phase process: one that works with the researcher to secure the data of the integrated database and one that focuses on retrieving, cleaning, and processing data so that it is ready for policy-driven research.

❖ **From Proposal to Publication: Chapin Hall's Research Procedures and Processes**

Like the other sites in the Actionable Intelligence for Social Policy network, Chapin Hall staff members understand that the data housed in the database do not belong to them. Rather, they are custodians of the data who ensure that the data are secure and that its use conforms to the guidelines that they have developed with the contributing agencies.

When researchers propose a study and request data, Chapin Hall staff ensures that the request reflects their mission. They then work with various researchers to develop a data-use plan to ensure that agency concerns regarding data use are met.

In some cases, researchers who have a long-standing relationship with Chapin Hall and the partnering agencies have full access to the data.

Dr. Robert Goerge generally serves as the point person for internal and external researchers and receives data requests and research proposals at least weekly. If the research proposal seems feasible and meets the aims of the integrated database, the next steps are taken, in which data needs and the procedures and processes regarding data use and practice are delineated. Chapin Hall relies on the University of Chicago's Institutional Review Board (IRB) to secure permissions for internal research. In addition, many of the agencies, such as the Department of Public Health and the Department of Children and Family Services, have formal IRBs that meet regularly to review and approve research proposals. Chapin Hall staff has outlined IRB procedures in their data sharing agreements to ensure that everyone understands the kinds of research studies that require IRB approval from their partner agencies. External researchers are responsible for securing IRB permissions with their own institutions and with the agencies that own the data.

Researchers generally contact Goerge directly, and once he has reviewed the proposal, he passes the proposal to his team who helps the researchers with the process of identifying and contacting the relevant agencies to ensure that the research meets the agency's needs and policy concerns. In some cases, researchers who have a long-standing relationship with Chapin Hall and its partnering agencies have full access to the data. For example, Joseph Doyle, the Alfred Henry and Jean Morrison Hayes Career Development Associate Professor of Economics at the Massachusetts Institute of Technology, has direct access to the data because he has secured permissions from the agencies that the IDS team works with and has been using the data for over a decade,

beginning when he was a graduate student. Approximately a dozen faculty members at the University of Chicago, DePaul University, University of Michigan, and the University of Illinois-Chicago have permission to access the data. These agreements are arranged depending on project-needs and data access is granted once the researchers have secured agreements with the necessary agencies. When the researchers complete their studies, they are expected to share their findings with the agencies that have supplied their data before they submit any publications. The agencies do not have a right to modify the research, but they can comment on the findings and make recommendations. Goerge said that the agencies are generally appreciative that they know the research findings before any publications are submitted and that there are rarely problems with this process either from external researchers or agency personnel.

In addition to meeting with agency staff to discuss research, Chapin Hall's integrated database staff members encourage researchers to publish their findings as short 3-4 page policy briefs on a publicly-accessible website. Even though peer-reviewed journals are the gold standard for university research, Goerge notes that the review process often takes too much time to have a direct impact on policy. By the time the journal reviews, accepts, and publishes the research, the policy makers and key stakeholders have usually already made a decision or moved on to consider other policy options. Publishing Chapin Hall's work in peer-reviewed journals is a secondary objective because of its mission to get the research out to those who need it as quickly as possible. Research that is a significant contribution to a field is submitted to peer reviewed journals. In addition, many of the policy makers and key stakeholders who are involved in the research do not have access to these publications since they are typically only

Goerge points out that “our mission is different from the university’s [mission]. Our mission is not to publish everything in peer-reviewed journals...it’s more to immediately inform policy makers of what we learn when we learn it” (Interview, Goerge, April 22, 2013).

available to researchers who are affiliated with university libraries. Goerge points out that “our mission is different from the university’s [mission]. Our mission is not to publish everything in peer-reviewed journals...it’s more to immediately inform policy makers of what we learn when we learn it” (Interview, Goerge, April 22, 2013).

Rather than submitting their findings to peer-reviewed journals, Chapin Hall staff members maintain a website where researchers and analysts publish research briefs that outline relevant literature, research methods, and study findings in a freely accessible manner. Over the past several years, the Illinois Integrated Database has provided data for researchers who are interested in examining the educational trajectories of Chicago youth from high school to college, identifying the amount of government social program participation overlap among Illinois families, and investigating whether third-grade reading level can be used as an indicator of potential performance on four future educational outcome measures (Goerge, Smithgall, Seshadri, & Ballard, 2010; Lesnick, Goerge, Smithgall, & Gwynne, 2010; Milesi, Lansing, Bell, Goerge, & Stagner, 2010). These studies are available on the publicly available website so that policy makers and analysts have timely access to them to inform policy decisions and future outcomes (Interview, Goerge, April 22, 2013).

❖ **Data Analytics Pipeline: Moving Data from the Contributing Agencies to the Integrated Database**

While Goerge oversees the procedures that move a research proposal to its final publication, Nila Barnes, the director of Chapin Hall’s Integrated Database on Child and Family Programs in Illinois, manages the data analytics pipeline, which includes data retrieval, cleaning, and processing so that the data are integrated and ready for policy-driven research. Chapin Hall’s data are maintained on two separate servers: one that supports identified data and another that support deidentified data. The server that houses the identified data has strict security protocols: it is not linked to the internet, is only available to staff when they are in Chapin Hall’s building, and is only accessible to staff with approved passwords and clearances. Usually, the contributing agencies transfer

data to Barnes using a secure file transfer protocol (FTP) or an encrypted DVD that only she can access. As Barnes said, each agency has its own way of processing the data. For example, the Chicago Public Schools maintain one data file while the Department of Child and Family Services often sends 12 or 13 files to Chapin Hall. While Chapin Hall often receives these files monthly, some agencies send their data quarterly or weekly. Like other sites, they act as a custodian of these data and only distribute the processed, de-identified files to researchers after the request has been made and a member of the Chapin Hall team has vetted and approved the research study and data requirements with its legal team and partnering agencies (Interview, Barnes, April 22, 2013).

Her role is “IT within the research arena” and combines the primary features of an IT specialist with research and business acumen.

In her interview, Nila Barnes noted important distinctions between the work that she does and the work that an individual in a typical information technologies (IT) position might do. IT workers, she argued, typically focus primarily on creating applications to support data use. Her role is “IT within the research arena” and combines the primary features of an IT specialist with research and business acumen. As the point person for data requests, Barnes must have a deep understanding of the latest IT technologies, the various research methodologies, and the integrated database holdings. When Chapin Hall first began to assemble its integrated database, analysts processed the data files, transferred the files securely to the researchers, and hoped that the researchers could make sense of the data files. The researchers would often inquire about the data file layout or data fields because these parameters were not clear. Barnes now works directly with the researchers to process Chapin Hall’s current data holdings in a custom-made way to ensure that researchers have the data they need to complete their work in a format that is user friendly and appropriate for their research design. This explains why her role is so critical to the integrated database’s functionality and data use for policy-driven research.

Typically, staff members rely on first and last name, birth dates, and social security numbers (when available) to link data and generate matches; they do not use race very often in their matches because racial categories are unreliable and are often different depending on the parameters that the agencies use to define race. They employ a probabilistic match system and clean the data in two ways. First, Barnes examines the data and cleans what she can by hand. Then, she relies on special programs that IT staff have written specifically for cleaning and reformatting purposes and are custom made for each dataset. Chapin Hall rarely excludes data during the data linkage process; rather, data that are undocumented are segregated from the rest of the data set so that researchers have a sense of the complete data set, the matched data set, and the undocumented data set. The data in Chapin Hall's integrated database are GIS-enabled and rely on ArcGIS software to determine address matches. One of the researchers handles this aspect of the data cleaning process to ensure that the data are clean and complies with ArcGIS standards (Interview, Barnes, April 22, 2013). Chapin Hall has held two security audits from external firms to ensure compliance with the latest data safety standards (Interview, Goerge, April 22, 2013).

Nila Barnes commented that it often takes many years for agencies to develop a trusting relationship with the IDS staff. This relationship is crucial to data exchange and data use agreements. In the past, the IDS staff members have leveraged particular projects to initiate conversations and nurture relationships with agencies that might have been reluctant to provide their data to Chapin Hall. For example, in 2001, the staff contacted Chicago Public Schools personnel about a study that they and the Illinois Department of Children and Family Services wanted to do and worked with the school district to ensure that data were protected and files were secure. Now, the Chicago Public Schools routinely provides data to Chapin Hall. This project sparked the conversations necessary to build a trusting relationship and to answer questions that the school staff might have about data management and analytic processes.

One of the most innovative aspects of Chapin Hall's data management and analytics processes is a unique tool, the Data Enclave, that Barnes has developed to track IDS research studies and data fields. This is a web application that allows policy makers

and researchers examine IDS holdings to understand the various data files that Chapin Hall maintains, the size of the files in their possession, and the number of records that exist over time. In addition, Enclave has a data dictionary where individuals who are interested in Chapin Hall's IDS can search the data elements that are in each database and see if data elements exist in other data sets. This feature provides policy makers and researchers with a way to understand how they might be able to link data using similar data elements across different data sets. These features enhance functionality. Researchers can either search across Chapin Hall's IDS holdings for a particular data element or they can search a specific file to determine the data elements within that data set. Finally, Enclave has a repository where Chapin Hall researchers can share their findings and work with others internally who might be interested in pursuing similar agendas.

When it is completed, Enclave will provide a central repository where researchers can learn about Chapin Hall's data holdings, share information about their work, and contribute to the policy-research knowledge base. The aim of this web-based portal is to offer an initial entry point where researchers can access the information that they might need to design their research studies through code sharing, meta data documentation, and finalized research reports. Enclave will also track the funding sources and data use for specific studies over time. This feature will provide Chapin Hall staff with a historical overview of the various studies that researchers have conducted using IDS data and the various funders who have supported this work. In her interview, Barnes stated that she designed and implemented this portal without additional funds and, thus, had to piece it together in stages using a variety of tools and technologies that she learned either from IT staff members or on her own. As she said, the project "literally grew organically" as she discussed with others how to organize the various information and learned how to create the back end supports necessary to enable this technology completely on her own. In the end, the Chapin Hall staff believe that creating Enclave was worth the investment as it will streamline the process for researchers who are interested in using the IDS data and help minimize the work that the staff has to do to support this work in the initial research design phase (Interview, Barnes, April 22, 2013).

❖ Economic and Political Realities for Sustaining Chapin Hall's Integrated Database

Chapin Hall's integrated database staff members generally rely on grant and contract funding—either from foundations, local, state or federal government—to acquire, clean, and archive IDS data. Chapin Hall also supports the IDS through its general fund when project funds fail to cover operating costs. However, institutional funds generally only cover approximately 5% of the operating costs. The rest of the funds are split across contracts (approximately 60%) and grants (approximately 40%). Although Chapin Hall has been associated with the University of Chicago for decades, as of July 1, 2013, it became an independent institution, although it is still affiliated with the University, and some staff, such as Goerge, continue to have University appointments. Chapin Hall's IDS research team will continue to run some grants through the University of Chicago, but they will be able to negotiate and execute contracts with government agencies and non-profit foundations through Chapin Hall directly. This new procedure will expedite these contracts in a more timely fashion that is free of the constraints that researchers typically encounter in large university bureaucracies.

Despite these advantages, Goerge admitted that developing continuity across budget cycles and finding funds to cover his operating expenses are the most challenging aspects of his position. He acknowledged that the lack of funding is clearly a constraint, but then said that this is also true for most university researchers. His main challenge stems from questions about how to charge researchers who rely on these data and use his staff to assist them with finding and accessing data to support their research. As he said, it is not clear whether Chapin Hall should charge external researchers for a percentage based on a full-time employee equivalency or a percentage rate based on total direct costs. It is a difficult decision because Goerge knows that the researchers could not afford to gather the kind of data that his team has assembled over several decades, but at the same time, many researchers do not understand this or do not have the funds to cover the actual expenses associated with gathering, cleaning, and organizing these data. As a result, Chapin Hall has oscillated among several different pricing structures. At one point, they charged researchers for IT time and IDS staff time.

As he says, without Chapin Hall’s policy research and integrated database, these agencies would “be less informed, and maybe would not have even developed their own capacities past where they are, because we gave them examples of what could be done.”

They have also asked researchers to calculate a percentage of the full time employee (FTE). Now, they are trying to derive a way to capture the actual time of the individuals who are involved in the project. However, this approach has several drawbacks because it spreads his staff across several different initiatives. For example, Nila Barnes, the database administrator, is currently covered by 14 different projects with varying degrees of commitment. Currently, staff members are engaged in several conversations about the best ways to deal with the challenges that they face in sustaining the work that they have done over the past 25 years (Interview, Goerge, April 22, 2013).

❖ **The Benefits of Chapin Hall’s Integrated Database for Policy-Driven Practice**

When one asks the Chapin Hall integrated database staff about the benefits of using integrated data for policy-driven research and practice, they agree that their work has the ability to impact and influence what state, county, and city agencies do. Goerge believes that this reflects Chapin Hall’s longstanding mission and serves as the core guiding principle for his staff. As he says, without Chapin Hall’s policy research and integrated database, these agencies would “be less informed, and maybe would not have even developed their own capacities past where they are, because we gave them examples of what could be done.” Even though the agencies did not act on all of the recommendations that his staff and researchers made, they did act on many research and policy findings (Goerge, Interview, April 22, 2013). In addition, all of the staff members commented about the work that they have done to sustain and maintain the

relationships and trust with the various agencies that contribute data to the database. Like many of the other sites in the AISP network, Chapin Hall staff members agree that these relationships have brought agencies together that worked in silos in the past. When the agencies see research studies and policy findings that draw on a variety of data, they begin to realize that the work that they are doing is actually connected in a variety of ways to the other agencies in the state, county, and city. Dr. Lucy Bilaver, a faculty member at Northern Illinois University's School of Nursing and Health Studies and Senior Researcher at Chapin Hall, suggests that Chapin Hall's integrated database research provides policy makers and analysts with a deeper understanding of the policies and practices that are effective as well as those that are not (Interview, Lucy Bilaver, April 22, 2013). This spurs conversations that might not have happened without the integrated database, which in turn, leads to better policies and outcomes for children and their families.

❖ About AISP

AISP is an initiative funded by the John D. and Catherine T. MacArthur Foundation through a grant to University of Pennsylvania Professors Dennis Culhane, School of Social Policy and Practice, and John Fantuzzo, Graduate School of Education. The principal aim of AISP is to improve the quality of education, health and human service agencies' policies and practices through the use of integrated data systems. Quality integrated data systems are designed to help executive leaders in municipal, county, and state government evaluate and establish effective programs for the people they serve.

❖ Works Cited—Data Use and Practices

Goerge, R. M., Smithgall, C., Seshadri, R., & Ballard, P. (2010). *Illinois Families and Their Use of Multiple Service Systems* (pp. 1–7). Chicago: Chapin Hall. Retrieved from <http://chapinhall.org/research/brief/illinois-families-and-their-use-multiple-service-systems>

Lesnick, J., Goerge, R. M., Smithgall, C., & Gwynne, J. (2010). *Reading on Grade Level in Third Grade: How Is It Related to High School Performance and College Enrollment?* (pp. 1–44). Chicago: Chapin Hall. Retrieved from <http://chapinhall.org/research/report/reading-grade-level-third-grade-how-it-related-high-school-performance-and-college-e>

Milesi, C., Lansing, J., Bell, K. C., Goerge, R. M., & Stagner, M. (2010). *Postsecondary Educational Trajectories of Urban Youth: Addressing Vulnerabilities and Barriers to Enrollment and Persistence* (pp. 1–8). Chicago: Chapin Hall. Retrieved from <http://chapinhall.org/research/brief/postsecondary-educational-trajectories-urban-youth-addressing-vulnerabilities-and-bar>

❖ Recommended Citation

Kitzmiller, Erika M. (2013). *IDS Case Study: Leveraging Chapin Hall's Mission to Enhance Child Well-Being*. Actionable Intelligence for Social Policy (AISP), University of Pennsylvania.