iScience

CellPress

Perspective

Managing Ideas, People, and Projects: Organizational Tools and Strategies for Researchers

Samuel Pascal Levin[1] and Michael Levin[2,*]

Primary Investigators at all levels of their career face a range of challenges related to optimizing their activity within the constraints of deadlines and productive research. These range from enhancing creative thought and keeping track of ideas to organizing and prioritizing the activity of the members of the group. Numerous tools now exist that facilitate the storage and retrieval of information necessary for running a laboratory to advance specific project goals within associated timelines. Here we discuss strategies and tools/software that, together or individually, can be used as is or adapted to any size scientific laboratory. Specific software products, suggested use cases, and examples are shown across the life cycle from idea to publication. Strategies for managing the organization of, and access to, digital information and planning structures can greatly facilitate the efficiency and impact of an active scientific enterprise. The principles and workflow described here are applicable to many different fields.

INTRODUCTION

Researchers, at all stages of their careers, are facing an ever-increasing deluge of information and deadlines. Additional difficulties arise when one is the Principal Investigator (PI) of those researchers: as group size and scope of inquiry increases, the challenges of managing people and projects and the interlocking timelines, finances, and information pertaining to those projects present a continuous challenge. In the immediate term, there are experiments to do, papers and grants to write, and presentations to construct, in addition to teaching and departmental duties. At the same time, however, the PI must make strategic decisions that will impact the future direction(s) of the laboratory and its personnel. The integration of deep creative thought together with the practical steps of implementing a research plan and running a laboratory on a day-to-day basis is one of the great challenges of the modern scientific enterprise. Especially difficult is the fact that attention needs to span many orders of scale, from decisions about which problems should be pursued by the group in the coming years and how to tackle those problems to putting out regular "fires" associated with the minutiae of managing people and limited resources toward the committed goals.

The planning of changes in research emphasis, hiring, grant-writing, etc. likewise occur over several different timescales. The optimization of resources and talent toward impactful goals requires the ability to organize, store, and rapidly access information that is integrated with project planning structures. Interestingly, unlike other fields such as business, there are few well-known, generally accepted guidelines for best practices available to researchers. Here we lay out a conceptual taxonomy of the life cycle of a project, from brainstorming ideas through to a final deliverable product. We recommend methods and software/tools to facilitate management of concurrent research activities across the timeline. The goal is to optimize the organization, storage, and access to the necessary information in each phase, and, crucially, to facilitate the interconnections between static information, action plans, and work product across all phases. We believe that the earlier in the career of a researcher such tools are implemented and customized, the more positive impact they will exert on the productivity of their enterprise.

This overview is intended for anyone who is conducting research or academic scholarship. It consists of a number of strategies and software recommendations that can be used together or independently (adapted to suit a given individual's or group's needs). Some of the specific software packages mentioned are only usable on Apple devices, but similar counterparts exist in the Windows and Linux ecosystems; these are indicated in Table 1 (definitions of special terms are given in Table 2). These strategies were developed (and have been continuously updated) over the last 20 years based on the experiences of the Levin group and those of various collaborators and other productive researchers. Although very specific software and platforms are indicated, to facilitate the immediate and practical adoption by researchers at all levels, the important thing is the strategies illustrated by the examples. As software and hardware inevitably change over the next few years, the fundamental principles can be readily adapted to newer products.

[1]Beverly, MA 01915, USA

[2]Allen Discovery Center at Tufts University, Suite 4600, 200 Boston Avenue, Medford, MA 02155-4243, USA

*Correspondence: michael.levin@tufts.edu

https://doi.org/10.1016/j.isci.2019.09.017

Name of Software	Purpose	Where to Purchase	Platform	Alternatives for Other Platforms
Adobe Acrobat	Document sharing and archival	Acrobat.adobe.com	OS X, Windows	Okular (Windows, OS X, Linux)
Box Sync	File backup and synchronization across devices	Box.com	OS X, Windows	Dropbox (Windows, OS X, Linux)
Carbon copy Cloner	Scheduled bootable backups of all or part of a drive	Bombich.com	OS X	Acronis True Image (Windows, OS X) AMANDA (Windows, OS X, Linux)
Crashplan Pro	Scheduled cloud backups across devices	Crashplan.com	OS X, Windows, Linux	Backblaze (Windows, OS X)
Calibre	Database of books	Calibre-ebook.com	OS X, Windows, Linux	NA
DevonThink	Document and information storage database	Devontechnologies.com	OS X	Microsoft OneNote (Windows, OS X) Zim (Windows, OS X, Linux)
Dropbox	File backup, storage, and synchronization between devices	Dropbox.com	OS X, Windows, Linux	Sync.com (Browser only, but will work on any OS)
EndNote	Automated management of references and creation of bibliographies in documents	Endnote.com	OS X, Windows	Zotero (Windows, OS X, Linux) JabRef (Windows, OS X)
Evernote	Document and information storage database	Evernote.com	OS X, Windows	NixNote (Windows, Linux) Notion (Windows, OS X)
MailSteward Pro	Long-term archival database for email	Mailsteward.com	OS X	Mailstore Server (Windows) Piler (Linux)
Microsoft Excel	Creation, management, and analysis of spreadsheet data	Products.office.com/excel	OS X, Windows	LibreOffice Calc (Windows, OS X, Linux) Apache Open Office Calc (Windows, OS X, Linux)
Microsoft Word	Creating and editing text documents	Products.office.com/word	OS X, Windows	Libre Office Writer (Windows, OS X, Linux) Apache Open Office Writer (Windows, OS X, Linux)
MindNode	Creating mind maps	Mindnode.com	OS X	Freemind (Windows, OS X, Linux) Mindomo (Windows, OS X, Linux, Browser)
OmniFocus	Organization and context-sensitive schedule of projects and plans	Omnigroup.com/omnifocus	OS X	RememberTheMilk (Windows, OS X, Linux) Asana (Browser-based, but a Windows client is available)
Spotlight	Title and content search for files in a file system	NA (it comes built-in with OS X and is not available on Linux or Windows)	OS X	Copernic Desktop Search (Windows) Albert (Linux) Cerebro (Windows, OS X, Linux)
PubCrawler	Automated search of PubMed databases for scientific papers	Pubcrawler.gen.tcd.ie	OS X, Windows, Linux	None found
Scrivener	Creating and editing of large project manuscripts	Literatureandlatte.com/scrivener	OS X, Windows	yWriter (Windows, OS X, Linux) Manuskript (Windows, OS X, Linux)
SuperDuper	Scheduled bootable backups of all or part of a drive	Shirt-pocket.com/SuperDuper	OS X	Acronis True Image (Windows, OS X) AMANDA (Windows, OS X, Linux)
Time Machine	Versioned, automated backups of files	NA (it comes built-in with OS X and is not available on Linux or Windows)	OS X	RollbackRx (Windows) Duplicati (Windows OS X, Linux)

Table 1. Software Packages and Alternatives

Term	Meaning
EPUB	A standardized format for digital books.
FTP	FTP stands for File Transfer Protocol. It is a protocol used to transfer files from one computer to another via a wired or wireless network.
Gantt chart	A type of bar chart used for project schedules, in which the tasks to be completed are shown as bars on the vertical axis, and time is shown on the horizontal axis, with the width of a given bar indicating the length of a given task. This facilitates planning by automating the tracking of milestone schedules and dependencies.
GTD	GTD stands for Getting Things Done. It is a productivity method created by productivity consultant David Allen that allows users to focus on those tasks that should be addressed in a given context and at the right timescale of planning, from current activities to life-long goals.
IP	IP stands for Intellectual Property, such as inventions and work products that are often patented or copyrighted.
Linux	Linux is a family of open-source operating systems created by Linus Torvalds in 1991, serving as an alternative to the commercial ones.
MTA	MTA stands for Materials Transfer Agreement—contracts that govern the transfer of research materials (e.g., DNA plasmids, cell lines) across institutions.
MySQL	MySQL is an open-source database management system, consisting of a server back end that houses the data and a front end that allows users to query the database in very flexible ways.
OCR	OCR stands for Optical Character Recognition—a process by which text is automatically recognized in an image, for example, converting a FAX or photo of a document into an editable text file.
PDF	PDF stands for Portable Document Format, which serves as a standard format for many different types of devices and operating systems to be able to display (and sometimes edit) documents.
PMID	PMID stands for PubMed ID—the unique identifier used in the PubMed database to refer to published papers.
SFTP	SFTP stands for SSH File Transfer Protocol but is often also referred to as Secure File Transfer Protocol. Its purpose is to transfer data over a network, similarly to FTP, but with added security (encryption).
SSH	SSH stands for Secure Shell. This allows a remote user to connect to the operating system of their computer via a terminal-like interface.
SSD	SSD stands for Solid State Drive. An SSD is a type of storage device for a computer that uses flash memory instead of a spinning disk, as in a typical hard drive. Compared with spinning hard drives, these are smaller, require less power, generate less heat, are less likely to break during routine use, and, crucially, enable vastly faster read and write speeds.
TB	TB stands for Terabyte—a unit of measuring file size on a computer. One terabyte is equivalent to one thousand gigabytes, one million megabytes, or one trillion bytes.
VNC	VNC stands for Virtual Network Computing—a desktop sharing system that transmits video signal and commands from one computer to another, allowing a user to interact with a remote computer the same way as if it were the computer they were currently using.
VPN	VPN stands for Virtual Private Network. A virtual private network allows connections to internet-based resources with high security (encryption of data).
WYSIWYG	WYSIWYG stands for What You See Is What You Get. This refers to applications where the output of text or other data being edited appears the same on-screen as it will when it is a finished project, such as a sheet of paper with formatted text (Microsoft Word and Scrivener are such, whereas LaTeX is not).
Windows	Windows refers to the operating system Microsoft Windows. It is one of the most common operating systems in use today and is compatible with the vast majority of applications and hardware.
XML	XML stands for Extensible Markup Language. Extensible Markup Language is a markup language used to encode documents such that they are readable by both humans and a variety of software.

Table 2. A Glossary of Special Terms

BASIC PRINCIPLES

Although there is a huge variety of different types of scientific enterprises, most of them contain one or more activities that can be roughly subsumed by the conceptual progression shown in Figure 1. This life cycle progresses from brainstorming and ideation through planning, execution of research, and then creation of work products. Each stage requires unique activities and tools, and it is crucial to

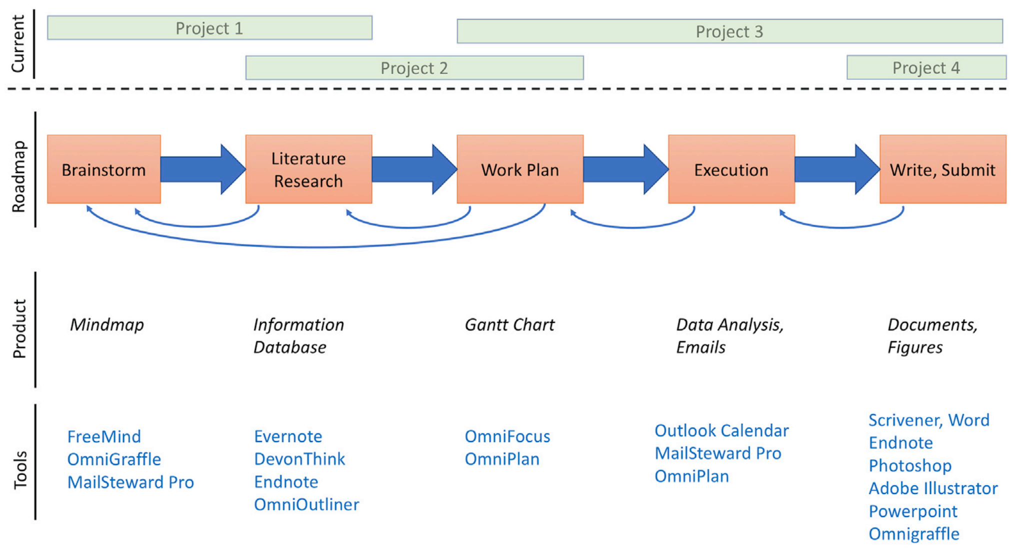

Figure 1. The Life Cycle of Research Activity

Various projects occupy different places along a typical timeline. The life cycle extends from creative ideation to gathering information, to formulating a plan, to the execution for the plan, and then to producing a work product such as a grant or paper based on the results. Many of these phases necessitate feedback to a prior phase, shown in thinner arrows (for example, information discovered during a literature search or attempts to formalize the work plan may require novel brainstorming). This diagram shows the product (end result) of each phase and typical tools used to accomplish them.

establish a pipeline and best practices that enable the results of each phase to effectively facilitate the next phase. All of the recommendations given below are designed to support the following basic principles:

- Information should be easy to find and access, so as to enable the user to have to remember as little as possible—this keeps the mind free to generate new, creative ideas. We believe that when people get comfortable with *not having to remember any details* and are completely secure in the knowledge that the information has been offloaded to a dependable system and will be there when they need it, a deeper, improved level of thinking can be achieved.
- Information should be both organized hierarchically (accessible by drill-down search through a rational structure) and searchable by keywords.
- Information should be reachable from anywhere in the world (but secure and access restricted). Choose software that includes a cell phone/tablet platform client.
- No information should ever be lost—the systems are such that additional information does not clog up or reduce efficiency of use and backup strategies ensure disaster robustness; therefore, it is possible to save everything.
- Software tools optimized for specific management tasks should be used; select those tools based on interoperability, features, and the ability to export into common formats (such as XML) in case it becomes expedient someday to switch to a newer product.
- One's digital world should be organized into several interlocking categories, which utilize different tools: activity (to-dos, projects, research goals) and knowledge (static information).
- One's activity should be hierarchically organized according to a temporal scale, ranging from immediate goals all the way to career achievement objectives and core mission.
- Storage of planning data should allow integration of plans with the information needed to implement them (using links to files and data in the various tools).
- There should be no stored paper—everything should be obtained and stored in a digital form (or immediately digitized, using one of the tools described later in this document).

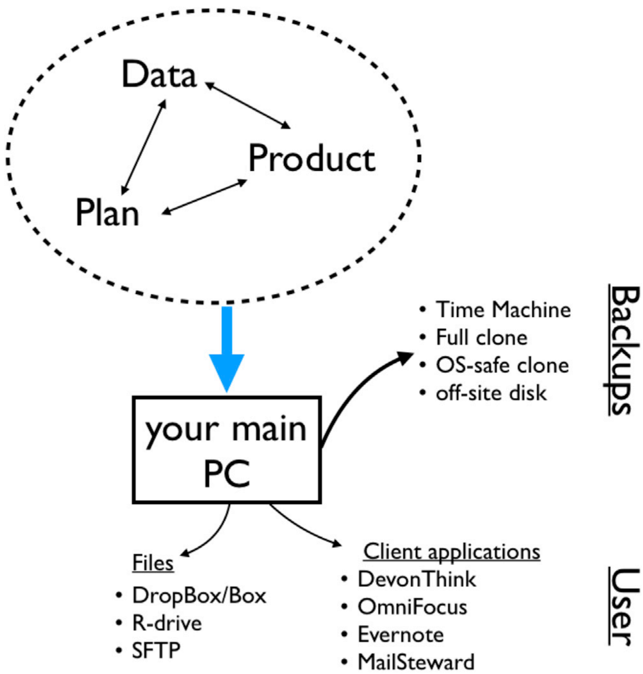

Figure 2. Schematic of Data Flow and Storage

Three types of information: data (facts and datasets), action plans (schedules and to-do lists), and work product (documents) all interact with each other in defining a region of work space for a given research project. All of this should be hosted on a single PC (personal computer). It is accessed by a set of regular backups of several types, as well as by the user who can interact with raw files through the file system or with organized data through a variety of client applications that organize information, schedules, and email. See Table 2 for definitions of special terms.

- The information management tasks described herein should not occupy so much time as to take away from actual research. When implemented correctly, they result in a net increase in productivity.

These basic principles can be used as the skeleton around which specific strategies and new software products can be deployed. Whenever possible, these can be implemented via external administration services (i.e., by a dedicated project manager or administrator inside the group), but this is not always compatible with budgetary constraints, in which case they can readily be deployed by each principal investigator. The PIs also have to decide whether they plan to suggest (or insist) that other people in the group also use these strategies, and perhaps monitor their execution. In our experience, it is most essential for anyone leading a complex project or several to adopt these methods (typically, a faculty member or senior staff scientist), whereas people tightly focused on one project and with limited concurrent tasks involving others (e.g., Ph.D. students) are not essential to move toward the entire system (although, for example, the backup systems should absolutely be ensured to be implemented among all knowledge workers in the group). The following are some of the methods that have proven most effective in our own experience.

INFORMATION TECHNOLOGY INFRASTRUCTURE

Several key elements should be pillars of your Information Technology (IT) infrastructure (Figure 2). You should be familiar enough with computer technology that you can implement these yourself, as it is rare for an institutional IT department to be able to offer this level of assistance. Your primary disk should be a large (currently, ~2TB) SSD drive or, better, a disk card (such as the 2TB SSD NVMe PCIe) for fast access and minimal waiting time. Your computer should be so fast that you spend no time (except in the case of calculations or data processing) waiting for anything—your typing and mouse movement should be the rate-limiting step. If you find yourself waiting for windows or files to open, obtain a better machine.

One key element is backups—redundant copies of your data. Disks fail—it is not a question of whether your laptop or hard drive will die, but when. Storage space is inexpensive and researchers' time is precious: team members should not tolerate time lost due to computer snafus. The backup and accessibility system should be such that data are immediately recoverable following any sort of disaster; it only has to be set up once, and it only takes one disaster to realize the value of paranoia about data. This extends also to laboratory inventory systems—it is useful to keep (and back up) lists of significant equipment and reagents in the laboratory, in case they are needed for the insurance process in case of loss or damage.

The main drive should be big enough to keep all key information (not primary laboratory data, such as images or video) in one volume—this is to facilitate cloning. You should have an extra internal drive (which can be a regular disk) of the same size or bigger. Use something like Carbon Copy Cloner or SuperDuper to set up a nightly clone operation. When the main disk fails (e.g., the night before a big grant is due), boot from the clone and your exact, functioning system is ready to go. For Macs, another internal drive set up as a Time Machine enables keeping versions of files as they change. You should also have an external drive, which is likewise a Time Machine or a clone: you can quickly unplug it and take it with you, if the laboratory has to be evacuated (fire alarm or chemical emergency) or if something happens to your computer and you need to use one elsewhere. Set a calendar reminder once a month to check that the Time Machine is accessible and can be searched and that your clone is actually updated and bootable. A Passport-type portable drive is ideal when traveling to conferences: if something happens to the laptop, you can boot a fresh (or borrowed) machine from the portable drive and continue working. For people who routinely install software or operating system updates, I also recommend getting one disk that is a clone of the entire system and applications and then set it to nightly clone the *data only*, leaving the operating system files unchanged. This guarantees that you have a usable system with the latest data files (useful in case an update or a new piece of software renders the system unstable or unbootable and it overwrites the regular clone before you notice the problem). Consider off-site storage. CrashPlan Pro is a reasonable choice for backing up laboratory data to the cloud. One solution for a single person's digital content is to have two extra external hard drives. One gets a clone of your office computer, and one is a clone of your home computer, and then you swap—bring the office one home and the home one to your office. Update them regularly, and keep them swapped, so that should a disaster strike one location, all of the data are available. Finally, pay careful attention (via timed reminders) to how your laboratory machines and your people's machines are being backed up; a lot of young researchers, especially those who have not been through a disaster yet, do not make backups. One solution is to have a system like CrashPlan Pro installed on everyone's machines to do automatic backup.

Another key element is accessibility of information. Everyone should be working on files (i.e., Microsoft Word documents) that are inside a Dropbox or Box folder; whatever you are working on this month, the files should be *inside* a folder synchronized by one of these services. That way, if anything happens to your machine, you can access your files from anywhere in the world. It is critical that whatever service is chosen, it is one that *synchronizes a local copy of the data that live on your local machine* (not simply keeps files in the cloud) —that way, you have what you need even if the internet is down or connectivity is poor. Tools that help connect to your resources while on the road include a VPN (especially useful for secure connections while traveling), SFTP (to transfer files; turn on the SFTP, not FTP, service on your office machine), and Remote Desktop (or VNC). All of these exist for cell phone or tablet devices, as well as for laptops, enabling access to anything from anywhere. All files (including scans of paper documents) should be processed by OCR (optical character recognition) software to render their contents searchable. This can be done in batch (on a schedule), by Adobe Acrobat's OCR function, which can be pointed to an entire folder of PDFs, for example, and left to run overnight. The result, especially with Apple's Spotlight feature, is that one can easily retrieve information that might be written inside a scanned document.

Here, we focus on work product and the thought process, not management of the raw data as it emerges from equipment and experimental apparatus. However, mention should be made of electronic laboratory notebooks (ELNs), which are becoming an important aspect of research. ELNs are a rapidly developing field, because they face a number of challenges. A laboratory that abandons paper notebooks entirely has to provide computer interfaces anywhere in the facility where data might be generated; having screens, keyboards, and mice at every microscope or other apparatus station, for example, can be expensive, and it is not trivial to find an ergonomically equivalent digital substitute for writing things down in a notebook as ideas or data appear. On the other hand, keeping both paper notebooks for immediate recording, and ELNs for organized official storage, raises problems of wasted effort during the (perhaps incomplete) transfer of information from paper to the digital version. ELNs are also an essential tool to prevent loss of institutional knowledge as team members move up to independent positions. ELN usage will evolve over time as input devices improve and best practices are developed to minimize the overhead of entering metadata. However, regardless of how primary data are acquired, the researcher will need specific strategies for transitioning experimental findings into research product in the context of a complex set of personal, institutional, and scientific goals and constraints.

FACILITATING CREATIVITY

The pipeline begins with ideas, which must be cultivated and then harnessed for subsequent implementation (Altshuller, 1984). This step consists of two components: identifying salient new information and arranging it in a way that facilitates novel ideas, associations, hypotheses, and strategic plans for making impact.

For the first step, we suggest an automated weekly PubCrawler search, which allows Boolean searches of the literature. Good searches to save include ones focusing on specific keywords of interest, as well as names of specific people whose work one wants to follow. The resulting weekly email of new papers matching specific criteria complements manual searches done via ISI's Web of Science, Google Scholar, and PubMed. The papers of interest should be immediately imported into a reference manager, such as Endnote, along with useful Keywords and text in the Notes field of each one that will facilitate locating them later. Additional tools include DevonAgent and DevonSphere, which enable smart searches of web and local resources, respectively.

Brainstorming can take place on paper or digitally (see later discussion). We have noticed that the rate of influx of new ideas is increased by habituating to never losing a new idea. This can be accomplished by establishing a voicemail contact in your cell phone leading to your own office voicemail (which allows voice recordings of idea fragments while driving or on the road, hands-free) and/or setting up Endnote or a similar server-synchronized application to record (and ideally transcribe) notes. It has been our experience that the more one records ideas arising in a non-work setting, the more often they will pop up automatically. For notes or schematics written on paper during dedicated brainstorming, one tool that ensures that nothing is lost is an electronic pen. For example, the Livescribe products are well integrated with Evernote and ensure that no matter where you are, anything you write down becomes captured in a form accessible from anywhere and are safe no matter what happens to the original notebook in which they were written.

Enhancing scientific thought, creative brainstorming, and strategic planning is facilitated by the creation of mind maps: visual representations of spatial structure of links between concepts, or the mapping of planned activity onto goals of different timescales. There are many available mind map software packages, including MindNode; their goal is to enable one to quickly set down relationships between concepts with a minimum of time spent on formatting. Examples are shown in Figures 3A and 3B. The process of creating these mind maps (which can then be put on one's website or discussed with the laboratory members) helps refine fuzzy thinking and clarifies the relationships between concepts or activities. Mind mappers are an excellent tool because their light, freeform nature allows unimpeded brainstorming and fluid changes of idea structure but at the same time forces one to explicitly test out specific arrangements of plans or ideas.

It is important to note that mind maps can serve a function beyond explicit organization. In a good mapped structure, one can look for symmetries (revealing relationships that are otherwise not obvious) between the concepts involved. An obvious geometric pattern with a missing link or node can help one think about what could possibly go there, and often identifies new relationships or items that had not been considered (Figure 3C), in much the same way that gaps in the periodic table of the elements helped identify novel elements.

ORGANIZING INFORMATION AND KNOWLEDGE

The input and output of the feedback process between brainstorming and literature mining is information. Static information not only consists of the facts, images, documents, and other material needed to support a train of thought but also includes anything needed to support the various projects and activities. It should be accessible in three ways, as it will be active during all phases of the work cycle. Files should be arranged on your disk in a logical hierarchical structure appropriate to the work. Everything should also be searchable and indexed by Spotlight. Finally, some information should be stored as entries in a data management system, like Evernote or DevonThink, which have convenient client applications that make the data accessible from any device.

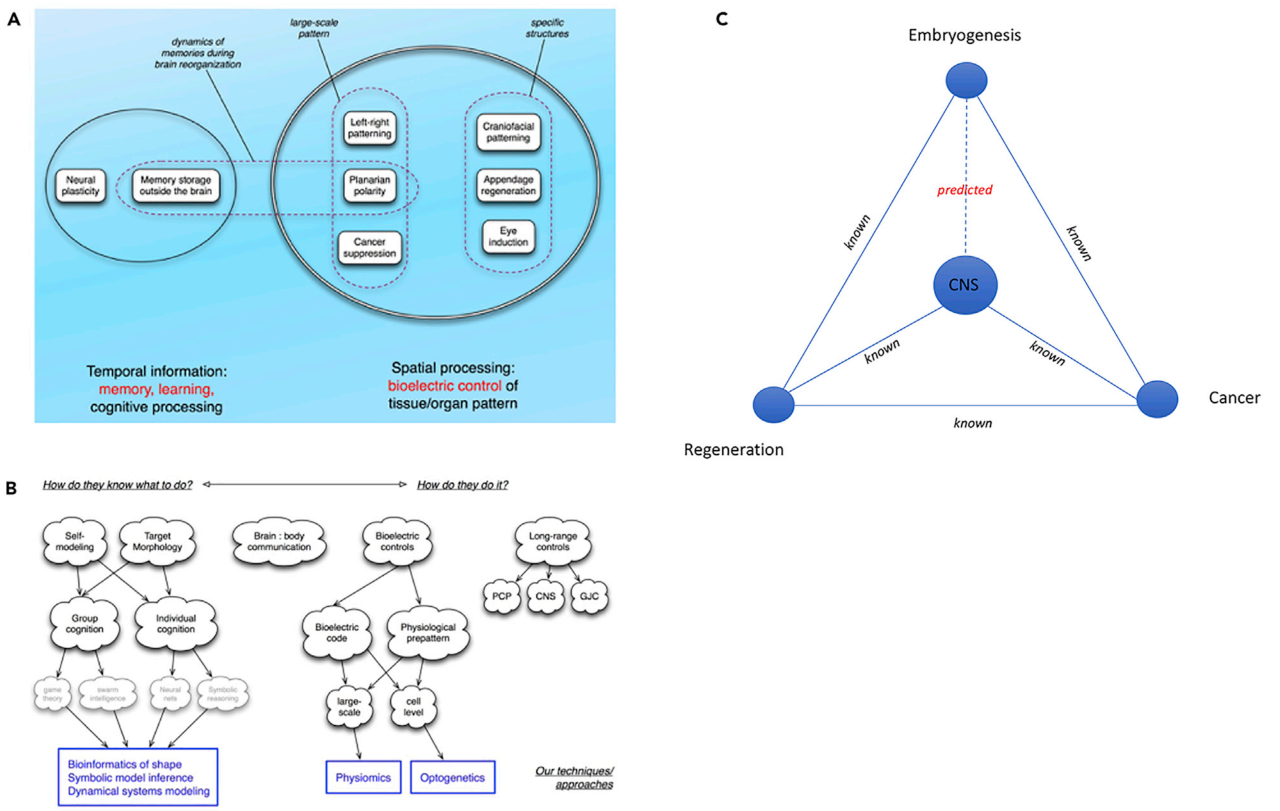

Figure 3. Mind Mapping

(A and B) The task of schematizing concepts and ideas spatially based on their hierarchical relationships with each other is a powerful technique for organizing the creative thought process. Examples include (A), which shows how the different projects in our laboratory relate to each other. Importantly, it can also reveal disbalances or gaps in coverage of specific topics, as well as help identify novel relationships between sub-projects by placing them on axes (B) or even identify novel hypotheses suggested by symmetry.

(C) Relationships between the central nervous system (CNS) and regeneration, cancer, and embryogenesis. The connecting lines in black show typical projects (relationships) already being pursued by our laboratory, and the lack of a project in the space between CNS and embryogenesis suggests a straightforward hypothesis and project to examine the role of the brain in embryonic patterning.

Notes in these systems should include useful lists and how-to's, including, for example:

- Names and addresses of experts for specific topics
- Emergency protocols for laboratory or animal habitats
- Common recipes/methods
- Lists and outlines of papers/grants on the docket
- Information on students, computers, courses, etc.
- Laboratory policies
- Materials and advice for students, new group members, etc.
- Lists of editors, and preferred media contacts
- Lists of Materials Transfer Agreements (MTAs), contract texts, info on IP
- Favorite questions for prospective laboratory members

Each note can have attachments, which include manuals, materials safety sheets, etc. DevonThink needs a little more setup but is more robust and also allows keeping the server on one's own machine (nothing gets uploaded to company servers, unlike with Evernote, which might be a factor for sensitive data). Scientific papers should be kept in a reference manager, whereas books (such as epub files and PDFs of books and manuscripts) can be stored in a Calibre library.

EMAIL: A DISTINCT KIND OF INFORMATION

A special case of static information is email, including especially informative and/or actionable emails from team members, external collaborators, reviewers, and funders. Because the influx of email is ever-increasing, it is important to (1) establish a good infrastructure for its management and (2) establish policies for responding to emails and using them to facilitate research. The first step is to ensure that one only sees useful emails, by training a good Bayesian spam filter such as SpamSieve. We suggest a triage system in which, at specific times of day (so that it does not interfere with other work), the Inbox is checked and each email is (1) forwarded to someone better suited to handling it, (2) responded quickly for urgent things that need a simple answer, or (3) started as a Draft email for those that require a thoughtful reply. Once a day or a couple of times per week, when circumstances permit focused thought, the Draft folder should be revisited and those emails answered. We suggest a "0 Inbox" policy whereby at the end of a day, the Inbox is basically empty, with everything either delegated, answered, or set to answer later.

We also suggest creating subfolders in the main account (keeping them on the mail server, not local to a computer, so that they can be searched and accessed from anywhere) as follows:

- Collaborators (emails stating what they are going to do or updating on recent status)
- Grants in play (emails from funding agencies confirming receipt)
- Papers in play (emails from journals confirming receipt)
- Waiting for information (emails from people for whom you are waiting for information)
- Waiting for miscellaneous (emails from people who you expect to do something)
- Waiting for reagents (emails from people confirming that they will be sending you a physical object)

Incoming emails belonging to those categories (for example, an email from an NIH program officer acknowledging a grant submission, a collaborator who emailed a plan of what they will do next, or someone who promised to answer a specific question) should be sorted from the Inbox to the relevant folder. Every couple of weeks (according to a calendar reminder), those folders should be checked, and those items that have since been dealt with can be saved to a Saved Messages folder archive, whereas those that remain can be Replied to as a reminder to prod the relevant person.

In addition, as most researchers now exchange a lot of information via email, the email trail preserves a record of relationships among colleagues and collaborators. It can be extremely useful, even years later, to be able to go back and see who said what to whom, what was the last conversation in a collaboration that stalled, who sent that special protocol or reagent and needs to be acknowledged, etc. It is imperative that you know where your email is being stored, by whom, and their policy on retention, storage space limits, search, backup, etc. Most university IT departments keep a mail server with limited storage space and will delete your old emails (even more so if you move institutions). One way to keep a permanent record with complete control is with an application called MailSteward Pro. This is a front-end client for a freely available MySQL server, which can run on any machine in your laboratory. It will import your mail and store unlimited quantities indefinitely. Unlike a mail server, this is a real database system and is not as susceptible to data corruption or loss as many other methods.

A suggested strategy is as follows. Keep every single email, sent and received. Every month (set a timed reminder), have MailSteward Pro import them into the MySQL database. Once a year, prune them from the mail server (or let IT do it on their own schedule). This allows rapid search (and then reply) from inside a mail client for anything that is less than one year old (most searches), but anything older can be found in the very versatile MailStewardPro Boolean search function. Over time, in addition to finding specific emails, this allows some informative data mining. Results of searches via MailStewardPro can be imported into Excel to, for example, identify the people with whom you most frequently communicate or make histograms of the frequency of specific keywords as a function of time throughout your career.

Figure 4. Scales of Activity Planning

Activities should be assigned to a level of planning with a temporal scale, based on how often the goals of that level get re-evaluated. This ranges from core values, which can span an entire career or lifetime, all the way to tactics that guide day-to-day activities. Each level should be re-evaluated at a reasonable time frame to ensure that its goals are still consistent with the bigger picture of the level(s) above it and to help re-define the plans for the levels below it.

With ideas, mind maps, and the necessary information in hand, one can consider what aspects of the current operations plan can be changed to incorporate plans for new, impactful activity.

ORGANIZING TASKS AND PLANNING

A very useful strategy involves breaking down everything according to the timescales of decision-making, such as in the Getting Things Done (GTD) philosophy (Figure 4) (Allen, 2015). Activities range from immediate (daily) tasks to intermediate goals all the way to career-scale (or life-long) mission statements. As with mind maps, being explicit about these categories not only force one to think hard about important aspects of their work, but also facilitate the transmission of this information to others on the team. The different categories are to be revisited and revised at different rates, according to their position on the hierarchy. This enables you to make sure that effort and resources are being spent according to priorities.

We also strongly recommend a yearly personal scientific retreat. This is not meant to be a vacation to "forget about work" but rather an opportunity for freedom from everyday minutiae to revisit, evaluate, and potentially revise future activity (priorities, action items) for the next few years. Every few years, take more time to re-map even higher levels on the pyramid hierarchy; consider what the group has been doing—do you like the intellectual space your group now occupies? Are your efforts having the kind of impact you realistically want to make? A formal diagram helps clarify the conceptual vision and identify gaps and opportunities. Once a correct level of activity has been identified, it is time to plan specific activities.

A very good tool for this purpose, which enables hierarchical storage of tasks and subtasks and their scheduling, is OmniFocus (Figure 5). OmniFocus also enables inclusion of files (or links to files or links to Evernote notes of information) together with each Action. It additionally allows each action to be marked as "Done" once it is complete, providing not only a current action plan but a history of every past activity. Another interesting aspect is the fact that one can link individual actions with specific contexts: visualizing the database from the perspective of contexts enables efficient focus of attention on those tasks that are relevant in a specific scenario. OmniFocus allows setting reminders for specific actions and can be used for adding a time component to the activity.

The best way to manage time relative to activity (and to manage the people responsible for each activity) is to construct Gantt charts (Figure 6), which can be used to plan out project timelines and help keep grant and contract deliverables on time. A critical feature is that it makes dependencies explicit, so that it is clear which items have to be solved/done before something else can be accomplished. Gantt charts are essential for complex, multi-person, and/or multi-step projects with strict deadlines (such as grant deliverables and progress reports). Software such as OmniPlanner can also be used to link resources (equipment, consumables, living material, etc.) with specific actions and timelines. Updating and evaluation of a Gantt chart for a specific project should take place on a time frame appropriate to the length of the next immediate phase; weekly or biweekly is typical.

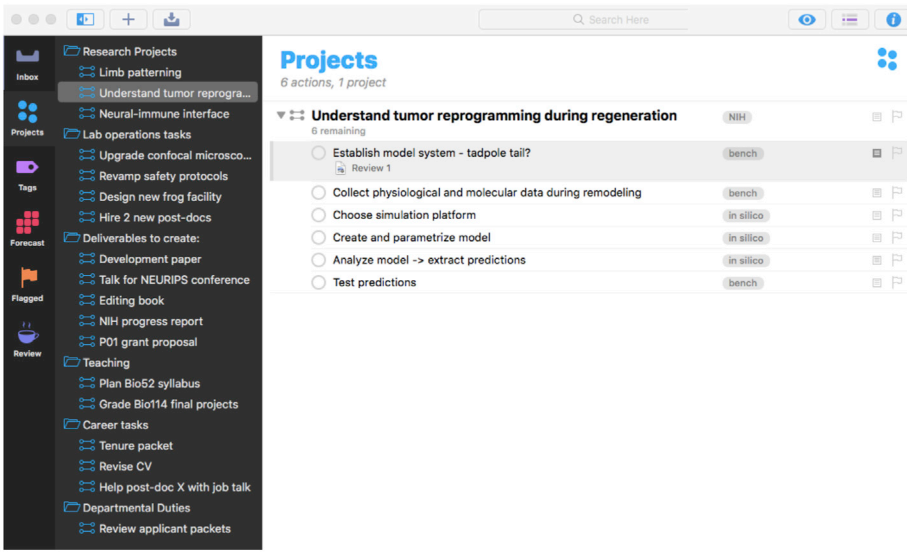

Figure 5. Project Planning

This figure shows a screenshot of the OmniFocus application, illustrating the nested hierarchy of projects and sub-projects, arranged into larger groups.

In addition to the comprehensive work plan in OmniFocus or similar, it is helpful to use a Calendar (which synchronizes to a server, such as Microsoft Office calendar with Exchange server). For yourself, make a task every day called ''Monday tasks,'' etc., which contains all the individual things to be accomplished (which do not warrant their own calendar reminder). First thing in the morning, one can take a look at the day's tasks to see what needs to be done. Whatever does not get done that day is to be copied onto another day's tasks. For each of the people on your team, make a timed reminder (weekly, for example, for those with whom you meet once a week) containing the immediate next steps for them to do and the next thing they are supposed to produce for your meeting. Have it with you when you meet, and give them a copy, updating the next occurrence as needed based on what was decided at the meeting to do next. This scheme makes it easy for you to remember precisely what needs to be covered in the discussion, serves as a record of the project and what you walked about with whom at any given day (which can be consulted years later, to reconstruct events if needed), and is useful to synchronize everyone on the same page (if the team member gets a copy of it after the meeting).

WRITING: THE WORK PRODUCTS

Writing, to disseminate results and analysis, is a central activity for scientists. One of the OmniFocus library's sections should contain lists of upcoming grants to write, primary papers that are being worked on, and reviews/hypothesis papers planned. Microsoft Word is the most popular tool for writing papers—its major advantage is compatibility with others, for collaborative manuscripts (its Track Changes feature is also very well implemented, enabling collaboration as a master document is passed from one co-author to another). But Scrivener should be seriously considered—it is an excellent tool that facilitates complex projects and documents because it enables WYSIWYG text editing in the context of a hierarchical structure, which allows you to simultaneously work on a detailed piece of text while seeing the whole outline of the project (Figure 7).

It is critical to learn to use a reference manager—there are numerous ones, including, for example, Endnote, which will make it much easier to collaborate with others on papers with many citations. One specific tip to make collaboration easier is to ask all of the co-authors to set the reference manager to use PMID Accession Number in the temporary citations in the text instead of the arbitrary record number it uses by default. That way, a document can have its bibliography formatted by any of the co-authors even if they have completely different libraries. Although some prefer collaborative editing of a Google Doc file, we have found a ''master document'' system useful, in which a file is passed around among collaborators by

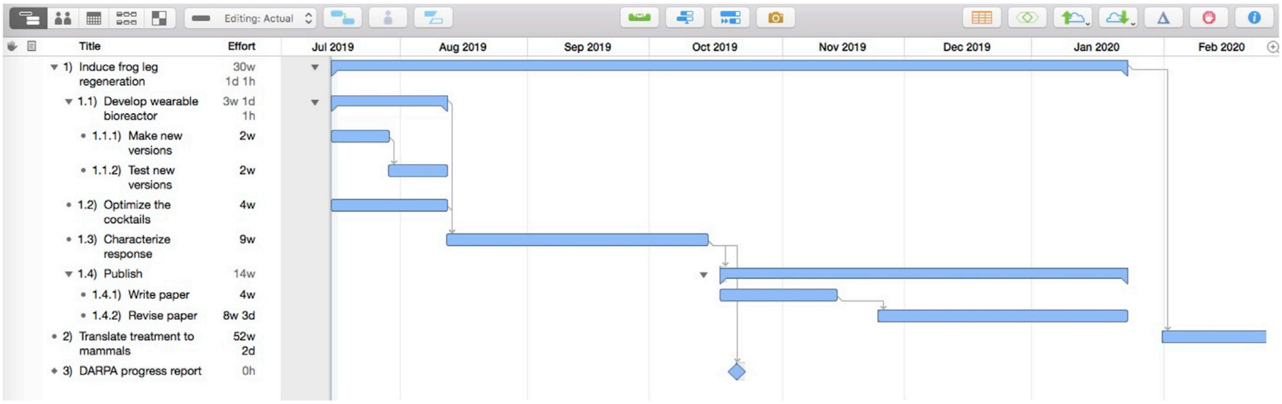

Figure 6. Timeline Planning

This figure shows a screenshot of a typical Gantt chart, in OmniPlan software, illustrating the timelines of different project steps, their dependencies, and specific milestones (such as a due date for a site visit or grant submission). Note that Gantt software automatically moves the end date for each item if its subtasks' timing changes, enabling one to see a dynamically correct up-to-date temporal map of the project that adjusts for the real-world contingencies of research.

email but only one can make (Tracked) edits at a time (i.e., one person has the master doc and everyone makes edits on top of that).

One task most scientists regularly undertake is writing reviews of a specific subfield (or Whitepapers). It is often difficult, when one has an assignment to write, to remember all of the important papers that were seen in the last few years that bear on the topic. One method to remedy this is to keep standing document files, one for each topic that one might plausibly want to cover and update them regularly. Whenever a good paper is found, immediately enter it into the reference manager (with good keywords) and put a sentence or two about its main point (with the citation) into the relevant document. Whenever you decide to write the review, you will already have a file with the necessary material that only remains to be organized, allowing you to focus on conceptual integration and not combing through literature.

Conclusion

The life cycle of research can be viewed through the lens of the tools used at different stages. First there are the conceptual ideas; many are interconnected, and a mind mapper is used to flesh out the structure of ideas, topics, and concepts; make it explicit; and share it within the team and with external collaborators. Then there is the knowledge—facts, data, documents, protocols, pieces of information that relate to the various concepts. Kept in a combination of Endnote (for papers), Evernote (for information fragments and lists), and file system files (for documents), everything is linked and cross-referenced to facilitate the projects. Activities are action items, based on the mind map, of what to do, who is doing what, and for which purpose/grant. OmniFocus stores the subtasks within tasks within goals for the PI and everyone in the laboratory. During meetings with team members, these lists and calendar entries are used to synchronize objectives with everyone and keep the activity optimized toward the next step goals. The product—discovery and synthesis—is embodied in publications via a word processor and reference manager. A calendar structure is used to manage the trajectory from idea to publication or grant.

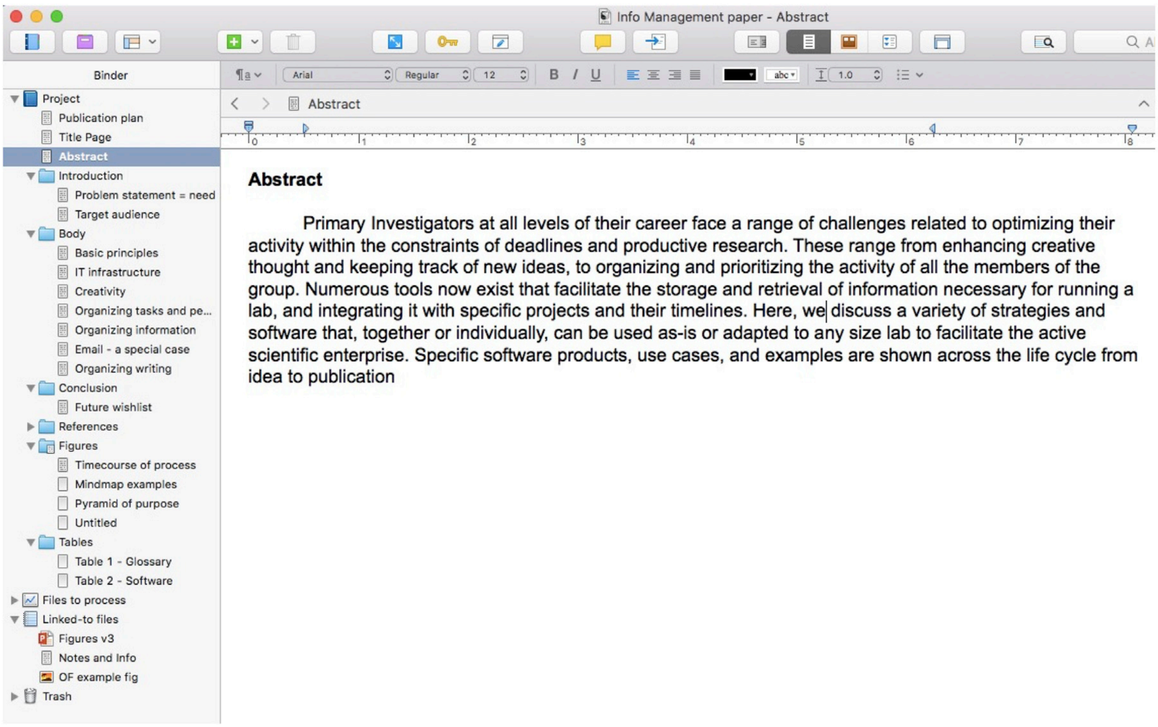

Figure 7. Writing Complex Materials

This figure shows a screenshot from the Scrivener software. The panel on the left facilitates logical and hierarchical organization of a complex writing project (by showing where in the overall structure any given text would fit), while the editing pane on the right allows the user to focus on writing a specific subsection without having to scroll through (but still being able to see) the major categories within which it must fit.

The tools are currently good enough to enable individual components in this pipeline. Because new tools are continuously developed and improved, we recommend a yearly overview and analysis of how well the tools are working (e.g., which component of the management plan takes the most time or is the most difficult to make invisible relative to the actual thinking and writing), coupled to a web search for new software and updated versions of existing programs within each of the categories discussed earlier.

A major opportunity exists for software companies in the creation of integrated new tools that provide all the tools in a single integrated system. In future years, a single platform will surely appear that will enable the user to visualize the same research structure from the perspective of an idea mind map, a schedule, a list of action items, or a knowledge system to be queried. Subsequent development may even include Artificial Intelligence tools for knowledge mining, to help the researcher extract novel relationships among the content. These will also need to dovetail with ELN platforms, to enable a more seamless integration of project management with primary data. These may eventually become part of the suite of tools being developed for improving larger group dynamics (e.g., Microsoft Teams). One challenge in such endeavors is ensuring the compatibility of formats and management procedures across groups and collaborators, which can be mitigated by explicitly discussing choice of software and process, at the beginning of any serious collaboration.

Regardless of the specific software products used, a researcher needs to put systems in place for managing information, plans, schedules, and work products. These digital objects need to be maximally accessible and backed up, to optimize productivity. A core principle is to have these systems be so robust and lightweight as to serve as an ''external brain'' (Menary, 2010)—to maximize creativity and deep thought by making sure all the details are recorded and available when needed. Although the above discussion focused on the needs of a single researcher (perhaps running a team), future work will address the unique needs of collaborative projects with more lateral interactions by significant numbers of participants.

ACKNOWLEDGMENTS

We thank Joshua Finkelstein for helpful comments on a draft of the manuscript. M.L. gratefully acknowledges support by an Allen Discovery Center award from the Paul G. Allen Frontiers Group (12171) and the Barton Family Foundation.

REFERENCES

Allen, D. (2015). Getting Things Done: The Art of Stress-free Productivity, Revised edition (Penguin Books).

Altshuller, G.S. (1984). Creativity as an Exact Science: The Theory of the Solution of Inventive Problems (Gordon and Breach Science Publishers).

Menary, R. (2010). The Extended Mind (MIT Press).