Как выбрать программное обеспечение для своего исследовательского проекта? Многие с безразличием относятся к выбору софта и работают с тем к чему уже привыкли, что уже установлено на рабочих и домашних компьютерах. В большинстве случаев мы полагаем, что никому не интересно, какими программами мы пользуемся, поэтому можно и не обращать на это внимание. Поскольку современная наука движется в сторону всё большей методической, фактической и процедурной прозрачности и воспроизводимости, скоро всё большему количеству исследователей придется объяснять не только то, как они, например, собирали данные и делали их анонимными, но и то, с помощью каких программ данные анализировали и визуализировали. Сейчас такое уже можно встретить в журналах, которые печатают материалы со статистическими расчетами, поскольку в ряде случаев разные программы будут давать разные результаты на одном и том же массиве данных. Понятно, что различия в большинстве случаев для неспециалиста покажутся несущественными, но наука — высокопрофессиональная деятельность и оценивают результаты нашей работы специалисты, а не профаны. Поэтому если вы работаете со статистикой, то необходимость выбора программного обеспечения для вас уже реальность. И даже если вы пользуетесь тем, что просто популярно и чем пользуются ваши коллеги, неосознанный выбор здесь уже присутствует.
И все же даже людям далеким от сбора и анализа различных данных ясно, что выбор одних программ будет выглядеть престижным, а других — не особенно.
Предложим «престижный» вариант:
— вы сами собрали массив данных в Сети,
— вы описали процедуру сбора и его результаты,
— вы сохранили в правильном формате, удобном для ваших коллег,
— вы проанализировали собранное при помощи свободного программного обеспечения,
— вы положили в свой собственный или организационный репозиторий.
В результате, вы сделали подборку данных, которую можете распространять свободно. Она полностью является результатом вашего интеллектуального труда. При этом любой человек с достаточными компетенциями сможет проверить точность ваших расчетов, поскольку программное обеспечение доступно для скачивания и использования без ограничений.
А вот менее «престижный» вариант:
— вы воспользовались базами, которые ограничивают распространение собранных в них материалов,
— вы считали в таких профессиональных программах, которые мало кто себе может позволить.
Конечно, наличие денег — тоже во многом признак успеха, но стоит признать, что прозрачность и воспроизводимость информации сейчас в научном мире имеют высочайший приоритет. Хорошая работа всегда найдет и читателя, и признание. Но в случае с закрытыми базами и платными программами торжество ваших идей будет доступно для проверки в меньшей степени.
Существует ряд неявных требований, которые было бы полезно учитывать каждому, кто планирует работать в научной среде. Далее последует небольшой список принципов, которые можно учитывать, если у вас есть желание грамотно выстроить работу долгосрочного проекта. Всё сказанное ниже, скорее всего, не особенно нужно для написания одной отдельной статьи, но если вы планируете продолжительную работу, то ознакомиться с принципами будет полезно. Процесс, хорошо продуманный в начале, всегда имеет более счастливое завершение.
Программное обеспечение — один из основных исследовательских инструментов. Ограничения программ становятся ограничениями исследовательской практики, а программные ошибки могут стать ошибками вашего исследования. Поскольку современный работник умственного труда проводит у монитора большую часть своего рабочего времени, стоит относиться к выбору программных решений предельно внимательно.
Принцип №1: нацеленность на решение ваших конкретных задач
Программное обеспечение подбирается не для того, чтобы быть модным, а чтобы решать конкретные исследовательские задачи, которые вы или ваш коллектив перед собой поставили (с учетом ваших финансовых возможностей). Предположим, что несмотря на то, что сейчас в моде R, если вы уже закупили Statistica или SPSS и привыкли работать с ними, то так и делайте. Если вы не закупили специализированый софт и хотите использовать простое программное обеспечение с открытым исходным кодом, то можете попробовать Gnumeric. Существует множество самых интересных решений, и в Сети можно найти разные профессиональные сайты, форумы, списки рассылок, где можно разобраться в достоинствах и недостатках различных программ. Если вы только начинаете свою исследовательскую карьеру, то не стоит стесняться спрашивать у коллег с большим опытом (если вы уже давно занимаетесь исследованиями, то вам и напоминать о возможности поинтересоваться мнением товарищей не нужно, ведь все так делают).
Принцип №2: соответствие профессиональным стандартам
Используемые вами программы должны быть не просто вам знакомы, удобны, но и соответствовать лучшим практикам, которые сложились в вашем исследовательском сообществе. Стоит ориентироваться на моду по двум причинам. Ваша исследовательская репутация состоит из многих элементов и использование передовых решений может хорошо дополнить ваши заслуги перед сообществом.
Самый простой способ понять, какие программы сейчас актуальны, пересмотреть публикации в ведущих изданиях по вашей тематике. Можно составить лист возможных программных решений, собрать информацию о цене, доступности для вашей инфраструктуры и потом принимать обоснованное решение.
Принцип №3: актуальность и доступ к поддержке
Программное обеспечение должно быть актуальным, то есть регулярно обновляться. Нередко улучшения, которые выходят, совершенно не заметны, а шансы столкнуться с теми ошибками, которые исправлены, и вовсе невелики, но в критический момент может произойти все что-угодно. Поэтому на сайтах всех производителей программ всегда пишут номер актуальной версии, чтобы пользователь мог понять, последняя у него версия или нет. Есть отдельные ситуации, когда обновление невозможно: например, если программа связана с устройством, которое больше не производится. В этих ситуациях нужно оценивать риски и прогнозировать дату, когда вы смените и устройство, и программы, придумывать, как будут переноситься данные и как нужно будет проводить исследования на новом оборудовании, чтобы результаты были совместимы со старыми данными.
Не менее важным, чем возможность обновлять свой софт, является доступ к поддержке. Если вы используете коммерческий продукт, то стоимость поддержки обычно закладывается в цену программы, поэтому вы можете законно рассчитывать на помощь и внимание в нужный момент. А если вы предпочитаете открытые решения, которые распространяются бесплатно, то уже не этапе выбора вам нужно оценить насколько активно и дружелюбно сообщество разработчиков и пользователей, которые теперь станут вашими помощниками по настройке и использованию программы. В любом случае, не забывайте, что даже если вы столкнулись с, казалось бы, неразрешимой проблемой, скорее всего, вы не первые и помощь можно найти.
Принцип №4: интероперабельность
Крайне важно, иметь возможность сохранять результаты вашей работы так, чтобы ими могли пользоваться те, у кого такой программы нет. Это и есть интероперабельность. Лучший вариант — использовать форматы файлов, которые соответствуют международным стандартам. Не во всех отраслях знания унификация происходит одинаково, например, в сфере библиографической информации все популярные пользовательские программы умеют передавать друг другу информацию. То же самое можно сказать о статистических программах. Но если вы занимаетесь качественными исследованиями, тогда все не так очевидно: начиная с того, в каком формате хранить видео и графику, заканчивая вопросом о том, какие способы разметки текста выбрать при расшифровке интервью.
В каждом отдельном случае разумно проверить, что делают коллеги и ориентироваться на наиболее продвинутые практики. При этом есть то, что можно порекомендовать всем: с большой опаской пользуйтесь программами, которые не позволяют выносить материалы вовне, поскольку если программа перестанет поддерживаться, вы не сможете больше работать со своим архивом. Лучше планировать все так, чтобы быть уверенным, что данные будут доступны и через 10-15 лет, на другой платформе даже если программу, в которой вы работаете, больше никто не обновляет. Самые популярные форматы с большой вероятностью не будут «забыты», но лучше не рисковать.
Осознанный подбор форматов для хранения пока встречается не часто. Например, несмотря на то, что известно, что предпочтительный форматом для хранения офисных документов являются ODT, ODS, ODP, а не DOCX, XLSX, PPTX (или их старые версии без X), массовой миграции на сертифицированные ISO форматы не произошло. Вероятно, не во всех ситуациях это необходимо и оправданно, но предполагать возможные будущие сложности стоит. Иначе придется разрабатывать программы для чтения устаревших форматов. Имея ввиду, что исследовательская культура движется в сторону большей открытости (прозрачности и воспроизводимости), требования к форматам хранения данных становятся более детально проработанными.
Принцип №5: финансовая доступность (не нужно ставить «ломанный» софт)
Программы, которые вы используете, должны быть легальными. Требование простое и понятное. Во многих случаях вы можете потребовать у работодателя, чтобы вам купили нужный софт. При этом скорее всего его даже приобретать не придется, потому что в Академии занимаются разными темами и, вероятно, что-то и для решения ваших задач уже купили. Стоимость таких вещей можно закладывать в сметы по грантам, хотя не все фонды позволяют делать такие закупки. Но планировать нужно.
Легальность программ — это не просто законодательное требование, но и вопрос безопасности. Если вы устанавливаете сломанный софт, то велик шанс того, что вместе с ним вы получаете вирус, который может красть какую-то значимую информацию профессионального или личного свойства, что безусловно нежелательно. Кроме того, сломанные программы часто не обновляются, а значит могут выполнять какие-то функции не так, как у других людей (см. принцип 3).
На данный момент нет нужды ставить ломанные программы, поскольку большинство задач можно решить бесплатными программами с открытым исходным кодом. Альтернативу уже знакомым программам можно посмотреть на сайте AlternativeTo.
Принцип №6: коллективная работа и инфраструктура
Если вы работаете в коллективе, а в большинстве крупных проектов на отдельных этапах это неизбежно, то нужно учитывать кто и какие функции будет выполнять, как вы будете контролировать процесс, с помощью чего рабочие материалы будут передаваться и редактироваться. Речь идет не только о совместном написании текстов, но и об удобном и безопасном хранении рабочих материалов.