PhD Manager

Technical specifications

PhD Manager is built on the Haplo open source platform.

The Haplo platform provides a flexible database tailored to storing information about the activities in complex organisations. It enables all types of information to be described and stored within a single integrated system. Sophisticated search enables information to be found, and automatic linking of related items supports serendipitous discovery of useful information.

The web-based user interface provides a familiar experience for users, reducing the need for training and shortening the time to embed the system into everyday use.

The platform includes support for collaboration and workflow, and the storage and processing of files in all major formats. It provides fine-grained permissions and full auditing. The high-levels of customisation are enabled by server-side JavaScript plugins.

Hostname (web address)

Haplo is hosted under your university domain name, with your choice of hostname, such as [research].youruniversity.ac.uk

The university will need to:

- supply a valid SHA-2 SSL certificate from a public CA, along with any intermediate certificates, ideally with a one-year expiry time. We can provide you with a CSR in any reasonable format.
- create a CNAME DNS record to point the chosen hostname to [university].infomanaged.co.uk

Branding

Your university's logo and branding colours are displayed throughout the system.

Data feed

Every university stores the information about their researchers and staff in different ways. We work with your IT team to get the data you need from the appropriate system and sent to Haplo. This enables Haplo to understand everyone's roles to route workflow to the right person.

Data is sent at an appropriate frequency, usually daily, by POSTing it to an https URL.

The likely data we would need from the university records system is:

* Name (title, first name, surname)
* Username
* Department
* Email
* Role (Researcher, Doctoral Researcher, Administrative staff)

* Job title (Staff only)
* Enrolment status
* Initial registration/enrolment date
* Intended award
* Visa type

While we prefer TSV (tab separate value) as an interchange format, we can use JSON, XML or any other structured text-based format.

If the university is able to provide all the required information in an LDAP directory, and allow Haplo to search it, we can use this instead of the data feed.

Authentication

Authentication is a separate mechanism to the user data feed to authenticate users when they log in. While Haplo has built-in authentication, we recommend integrating with the university's identity server for single sign on.

Haplo currently supports LDAP (including ActiveDirectory) and OAuth2 for Google Apps, but we can add support for extra identity providers, and implement custom login user interface.

If LDAP is chosen, we need the university to:

- Create an LDAP service account: Provide a hostname, username and password for your LDAP server. We only support encrypted connections.
- Create firewall rules for LDAP server: Allow access from our hosted networks (we will provide you with a list of network blocks which may make requests)
- LDAP certificates: Provide details of all SSL certificates which may be presented by the LDAP server (if they do not necessarily have the same hostname as the one we use to connect)
- LDAP search path: Provide details of which parts of the tree we should search for users, and which attribute is used for the username. (We check passwords by searching for the username in the tree to find the user's full DN, then attempt another connection with that DN and the provided password.)

Email notifications

Haplo will send email notifications to users on behalf of the university. There are a number of options for how we provide this, the two main options are 1) the university set up DNS SPF records, etc, to alert external spam filters that we're a valid source of email, and whitelisting within your email servers, or 2) we send emails to you via SMTP, which you can then deliver to the recipient.

Hosting and security

Haplo is offered as a hosted solution, using a fully redundant internet connection, with an 99.9% SLA from our internet provider, at a co-location facility in the UK. There is no single point of failure in the networking: Each server has two physical connections to the network with automatic failover using

probe-based liveness testing. These physical connections connect to different switches, which connect in a cross-over pair to a redundant firewall cluster, which in turn connects to two network backbone routers, which each have multiple redundant paths to the internet at large. Our network has had 99.995% availability over the last 4 years.

We have the ability to utilise private peering arrangements with other networks, if required and contractually feasible. Our network provider peers with JANET.

Servers

As a matter of policy, we own all the server hardware so we can offer strict data protection assurances to our customers. Only Haplo employees have access to our servers.
All server components with moving parts are in hot-swappable redundant pairs (power supplies, fans, storage) and monitored by an integrated service processor. The facility has two redundant power feeds, with backup generators. The servers can operate with only one power supply connected, and each feed has enough capacity to power the entire datacentre.

Our servers have had 99.995% availability over the last 4 years.

Data integrity

All storage uses mirrored disks for redundancy (equivalent of RAID 1). Future bulk storage may use other RAID levels for space efficiency.

All data is stored on ZFS filesystems with full cryptographic checksumming, for assurance that any hardware errors will be detected and corrected. Daily snapshots are taken, and every week, every bit of data on the disks are checked for proactive detection of hardware faults. In addition to the filesystem level checksumming, files managed by the platform are checksummed at the application level, and those checksums are verified.

Security and availability

Our firewalls only allow http and https traffic from the internet at large. Administrative access is controlled by certificates (not passwords) and requires VPN access. Our deployment system uses cryptographically signed software archives, ensuring that only authorised software is installed.

Each server and Platform instance is independent (no single point of failure within the cluster as a whole), and all traffic is encrypted. The servers are physically located in a locked rack in the datacentre. Access to the rack and datacentre floor is controlled by 24/7 manned security, and all visits are escorted and logged.

Monitoring

We use internal monitoring of server health, and external monitoring of network availability.

Backup & disaster recovery

We maintain equipment in a second datacentre. The live and backup servers mutually authenticate using strong cryptography. Data is transferred over encrypted connections between data centres using the public Internet.

Our current time to restore from a backup is under 4 hours. In a disaster recovery scenario, we can build a new server, automatically, in under 15 minutes starting from taking a new server out of the box. To ensure that our backup and disaster recovery mechanisms always work, we use the same processes for our day to day administration. For example, the normal software deployment and upgrade process uses exactly the same mechanism as would be used for disaster recovery.

To verify the backup system is working as designed, test restores are made on a regular basis to test the ability of the system to restore a backup, using Customer data chosen at random.

30 days of complete backups are kept. Daily snapshots are taken and retained for 14 days. One in 5 daily snapshots is retained for 90 days. Backups older than 90 days are deleted.

Archive service

We provide an optional archive service where a copy of a client's data is made available to them to maintain their own independent backup. The archive service is intended as a one-off or less regular service and a small charge is made for each archive operation.

Browser support

To access Haplo, users need to use Web browsers which support the SSL SNI extension. All modern web browsers support SNI.

Haplo can be used on both PCs and Macs, using browsers including Internet Explorer, Microsoft Edge, Firefox, Google Chrome and Safari. It can be used on mobile devices.

We will support any modern standards compliant web browser, and commit to supporting all the browsers supported by Microsoft Office 365 and Google Apps for Education.

Open source

The underlying Haplo platform is open source, released under the Mozilla Public License v2. It's written in Java, JRuby with applications written in JavaScript.

The platform code and documentation are available at http://haplo.org

We are working on open sourcing most of the core functionality of our Higher Education products. Features specific to a single client will not be open sourced for reasons of client confidentiality, but will be made available to the relevant client.

Connections to other systems

Haplo provides an extensive API to enable data to be fed to and from other systems, and custom APIs can be developed as required.

Permissions

Access permissions will be applied according to the requirements of the institution. Most permissions will be administered in the normal user interface by authorised users, and not require system administration privileges.

Key users can be granted the ability to 'impersonate' any other user, enabling them to see and take action by the other user for providing support.

Audit

A full-audit trail is maintained to show which actions were taken by which user, taking impersonation into account. The audit trail audits all logins, updates and deletions to records.

PhD Manager

Service levels

Availability target	
Service availability (not including planned maintenance)	99.5%
Planned maintenance	Up to 1 hour per month

Availability (% of calendar month)	Reimbursement rate (% of monthly Charges)	Maximum Downtime (hours per calendar month)
99.5% and above	0	3.6
99% - 99.49%	5	7.2
95% - 98.9%	10	36
90% - 94.9%	15	72
less than 90%	20	More than 72 hours

The Haplo team will notify you by email at least 5 working days in advance of scheduled downtime required for planned maintenance which is expected to exceed 30 seconds or take place within normal UK business hours. Where Planned maintenance is not expected to exceed 30 seconds Downtime, it will occur at the times listed in the Planned maintenance schedule.

Planned maintenance schedule for maintenance not exceeding 30 seconds	
Monday - Friday, excluding Bank Holidays	8am, 7pm (UK)
Saturday - Sunday, and Bank Holidays	All day

PhD Manager

The Haplo team will notify you of any report of non-scheduled downtime, and investigate and remedy it using suitably qualified personnel in line with the published response and fix time targets.

Response and fix time targets	Core service hours		Non-core service hours	
Severity	Response	Fix	Response	Fix
1. Critical The problem severely impacts your use of the software in a production environment (such as loss of data or your production systems not functioning.) The situation halts your business operations and no procedural workaround exists.	1 hour	Continuous effort	2 hours	Continuous effort
2. High The software is functioning but your use in a production environment is severely reduced. The situation is causing a high impact to portions of your business operations and no procedural workaround exists.	2 hours	1 business day	2 core service hours	1 business day
3. Medium The problem involves partial, non-critical loss of use of the software in a production environment or development environment. For production environments, there is a medium-to-low impact on your business, but your business continues to function, including by using a procedural workaround. For development environments, where the situation is causing your project to no longer continue or migrate into production.	4 hours	4 business days or as agreed	4 core service hours	4 business days or as agreed
4. Low A general usage question, reporting of a documentation error, or recommendation for a future product enhancement or modification. For production environments, there is low-to-no impact on your business or the performance or functionality of your system. For development environments, there is a medium-to-low impact on your business, but your business continues to function, including by using a procedural workaround.	6 hours	8 days or as agreed subject to product roadmap schedule	6 core service hours	8 days or as agreed subject to product roadmap schedule

Service	Days	Time
Core service hours	Monday - Friday (excluding Bank Holidays)	09.00 - 17.30 (UK)
Non-core service hours	Monday - Friday Saturday - Sunday Bank Holidays	17.31 - 08.59 (UK) All day All day

PhD Manager

About Haplo

Haplo Services is an innovative British company, combining Information Management and Information Technology expertise to deliver tailored solutions to complex information management problems.

The Haplo platform has been in commercial use continuously since November 2007. The platform was open sourced early in 2015.

Haplo Services have built an enviable reputation for the quality of its solutions, a relentless focus on information security, and high-levels of expert customer service.

Established in 2007, Haplo is based in London.

hello@haplo-services.com

Unit C 1st floor, Emperor House

Dragonfly Place, London SE4 2FL

020 7047 1111