Compound Passport Service: supporting corporate collection owners in open innovation

David M. Andrews[1], Sébastien L. Degorce[1], David J. Drake[2], Magnus Gustafsson[3], Kevin M. Higgins[2] and Jon J. Winter[1]

[1] Oncology iMed Chemistry, AstraZeneca, Mereside, Alderley Park, Macclesfield SK10 4TF, UK
[2] R&D Information, AstraZeneca, Mereside, Alderley Park, Macclesfield SK10 4TF, UK
[3] R&D Information, AstraZeneca, Pepparedsleden 1, 431 83 Mölndal, Sweden

A growing number of early discovery collaborative agreements are being put in place between large pharma companies and partners in which the rights for assets can reside with a partner, exclusively or jointly. Our corporate screening collection, like many others, was built on the premise that compounds generated in-house and not the subject of paper or patent disclosure were proprietary to the company. Collaborative screening arrangements and medicinal chemistry now make the origin, ownership rights and usage of compounds difficult to determine and manage. The Compound Passport Service is a dynamic database, managed and accessed through a set of reusable services that borrows from social media concepts to allow sample owners to take control of their samples in a much more active way.

Introduction

The challenges of discovering and developing novel therapeutics have been well documented [1]; and the combination of the 'low-hanging fruit' of drug targets having been picked off [2] along with the challenge to maintain the pace of new discovery has led to an increase in the complexity of targets and disease pathways in discovery portfolios. Additionally, the pharmaceutical industry has realigned resources away from early R&D [3], making industry more reliant on collaboration with academic groups to share the risks (and rewards) of conducting discovery and early validation efforts [4].

These efforts are frequently captured under the generic term 'open innovation', first coined by Henry Chesbrough in 2003 [5]. Since then, a huge variety of definitions for open innovation have been suggested; the authors prefer the definition adopted by the Wellcome Trust: 'The process of innovating with others for shared risk and reward to produce mutual benefits for each organisation, creating new products, processes or ideas that could not otherwise have been achieved alone, or enabling them to be achieved more quickly, cheaply or efficiently' [6].

Corresponding author: Andrews, D.M. (david.andrews@astrazeneca.com)

We found that in AstraZeneca (AZ), as is the case across the pharmaceutical industry [7], many more collaborative agreements were being put in place in which the rights for assets could reside with a partner, exclusively or jointly. With more opportunities being investigated to take advantage of our in-house assets, we needed to improve our ability to ensure that compounds subject to a contractual agreement with third parties are managed and used in accordance with AZ's obligations. Such agreements could mean compounds should be restricted to agreed tests and/or prevented from being shared with other parties.

The problem: compound rights tracking

Many corporate screening collections have been built on the premise that collection members that had been generated in-house and were not the subject of paper or patent disclosure were proprietary to the company. Collaborative screening arrangements [8] and medicinal chemistry [9,10] now make the origin, ownership rights and contractually governed usage of compounds difficult to determine and complex to manage. When we looked at the compound management tools available within our own organisation (or those available from vendors) we found that, in general, solutions were monolithic one-size-fits-all packages and

http://dx.doi.org/10.1016/j.drudis.2015.06.011

lacked the information granularity necessary to answer key questions around compound-asset rights: compound and sample restrictions were either single sample or all examples from a project; delegation of approval was difficult and all approval was manual; approval tracking and compliance monitoring was difficult and error-prone; in consequence it was difficult to provide partners with a record of requests for 'their' samples.

The root cause was that the design of the compound asset infrastructure predated the emergence of shared risk, collaborative, open innovation projects and the infrastructure had been designed against a background where there was a presumption that a company could solely own the rights to the portion of its compound library that was not in the public domain. Together, this created the risk that AZ scientists could unwittingly release the structures of early-stage hits to collaborators that were already the subject of an agreement with another collaborator. Put simply, the infrastructure of material transfer agreements and confidential disclosure agreements was very good at tracking the supply of single compounds but could not cope with the tens to thousands of compounds that needed to be correctly tracked as a result of open innovation collaborations. We aimed to design a dynamic database, managed and accessed through a set of reusable services that borrowed from social media concepts to allow sample owners to take control of their samples in a much more active way. Our design is driven by the vision that a greater number of collaborative agreements are being put in place and that, within those agreements, compound rights can be shared or reside with AZ or the collaboration partner. In turn this drove the need to improve our ability to keep our contractual agreements and progress compounds through approval flows in a timely and efficient manner. In addition to making it easy to record and update details of collaborations, we wanted simply and quickly to add, edit and remove compound assets, as well as to provide fast, reliable and automated approval where possible or to alert approvers where a manual approval is required. Finally, we wanted scientists to be able to determine compound status easily; able to request approval to take specific actions (e.g. testing) within the context of a system that maintained a permanent record.

For the service to function correctly, for each open innovation compound, three pieces of information (metadata) need to be captured and kept up-to-date: (i) who owns any rights to the compound – AZ, the partner or are the rights shared? (ii) Can the compound be tested freely within AZ or does the collaboration agreement indicate that a collaboration coordinator needs to approve test requests? (iii) Has the compound been provided to the partner structure-blinded or has the compound structure been shared with the partner? The service can control 'trafficking' as well as maintain a permanent compound 'life history'. It can be interrogated and receive updates via calls from other systems. Hence, we refer to it as the Compound Passport Service (CPS). Our shared vision and understanding of the underlying metadata allowed us to formulate the main concepts of the system and associated 'use cases'.

System concepts and use cases

CPS centres on assets, which are entities reflecting agreed constraints of compound usage by third parties and are grouped together in asset rights groups (ARG). Additionally, an ARG is a group of rights and rules applied to a number of assets. For each ARG a number of roles can be defined: coordinator – sets up and manages the ARG; delegate – a deputy for the coordinator who can add or remove approvers and add or update assets; approver – a person able to approve or reject requests manually, when the system is unable to make an automatic decision. For each ARG it is also possible to delineate request rules that define which requests can be automatically approved or rejected (e.g. that all compounds within an ARG can be tested in a specific test without any manual approval needed). A user in another system that is fully integrated with CPS becomes a requester when asking for approval to perform a specific action, for example to run a test on a specific compound and CPS responds with the following: 'approved', 'rejected' or 'manual approval needed' (Fig. 1).

This structure provides a framework that allows users to take control of their compounds in real-time and in a very granular way. It also has the potential to speed up the flow of compounds through the design-make-test-analyse (DMTA) cycle [11] by giving users the option to set up rules for automatic approval or rejection of tests. To enable these goals, we considered seven critical scenarios (use cases) that the system needed to service (Table 1).

More-efficient approval flows and search

The CPS was designed to manage complex compound sharing rules and requirements over the entire lifecycle of a collaboration project. The following is a case history of a recent project that was used to help design a system with the flexibility required. A collaboration project had two chemical series: A and B. Series A originated from screening of the AZ compound collection, and samples were tested by the partner organisation in a structure-blind format. Synthetic optimisation was performed by AZ but the compounds were never judged to be of sufficient selectivity to merit sharing with the partner. Owing to their origin in a collaboration project, however, series A compounds were not permitted to be tested outside the originating project. Later, the compounds were of no further interest to the project and permitted to be used by AZ projects for any purpose.

Series B originated from the partner organisation, and samples were initially shared for testing by AZ in a structure-blind format. After some months of optimisation, the chemical structures were shared with AZ but the ownership of the compounds was retained by the partner. Later still, the series was judged of sufficient quality for the intellectual property to be shared jointly between the two organisations. Finally, after the biological target hypothesis was devalidated, notional ownership of the compounds was returned to the partner organisation and no further testing was permitted by AZ (Fig. 2).

All of the compound status changes in this scenario are mapped on to three key properties that are captured by the CPS (Table 2). The owner of the ARG is able to change the properties of individual or groups of compounds as required, independently, as the project evolves. Such changes are recorded with time stamps as 'transitions', and can be tracked over the lifetime of an ARG. The ability to track all such transitions increases the transparency of compound ownership to AZ and partner organisations, prevents unauthorised testing of samples by other project teams within AZ and enables questions of the provenance of compounds to be easily resolved.

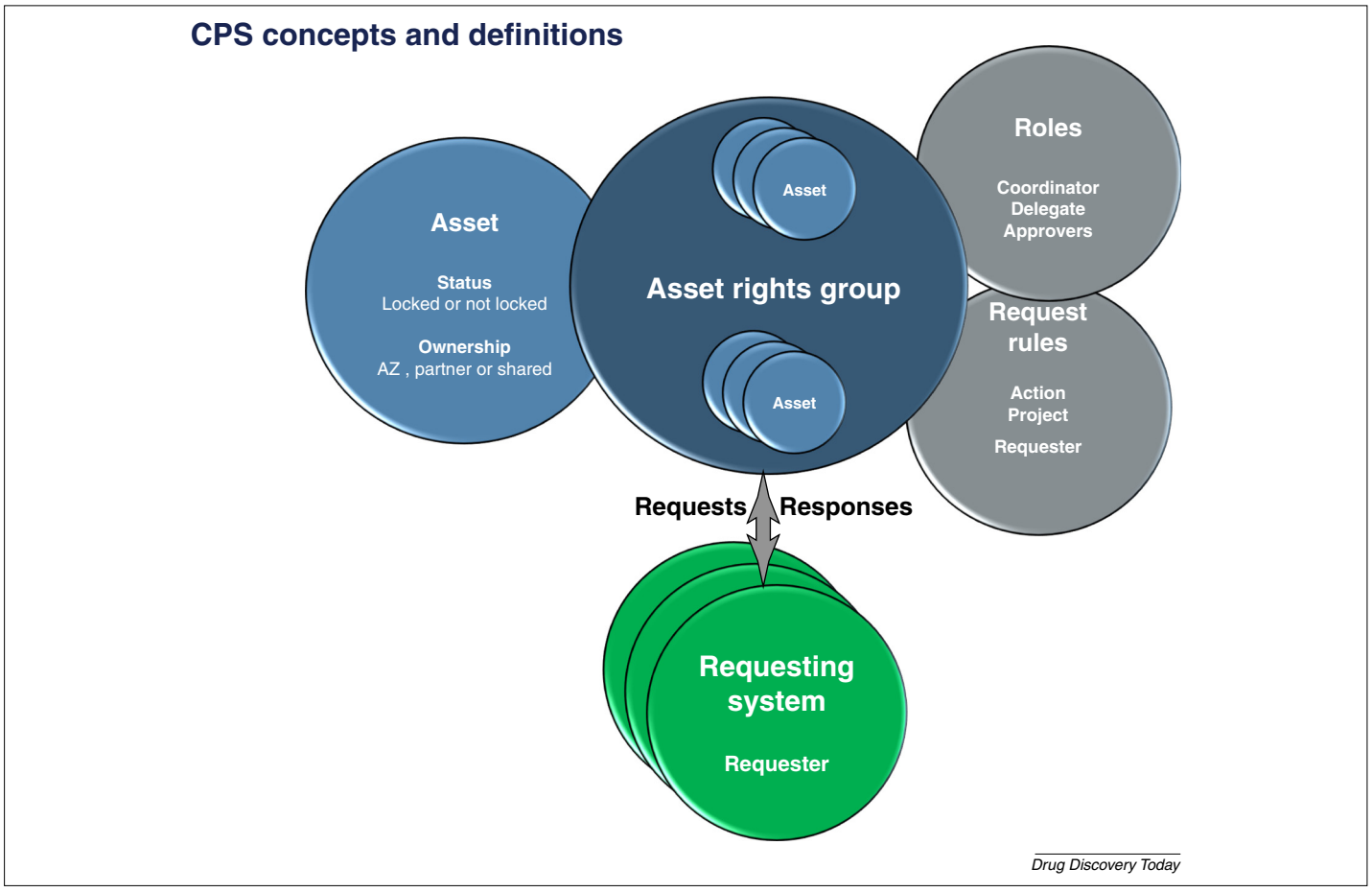


FIGURE 1

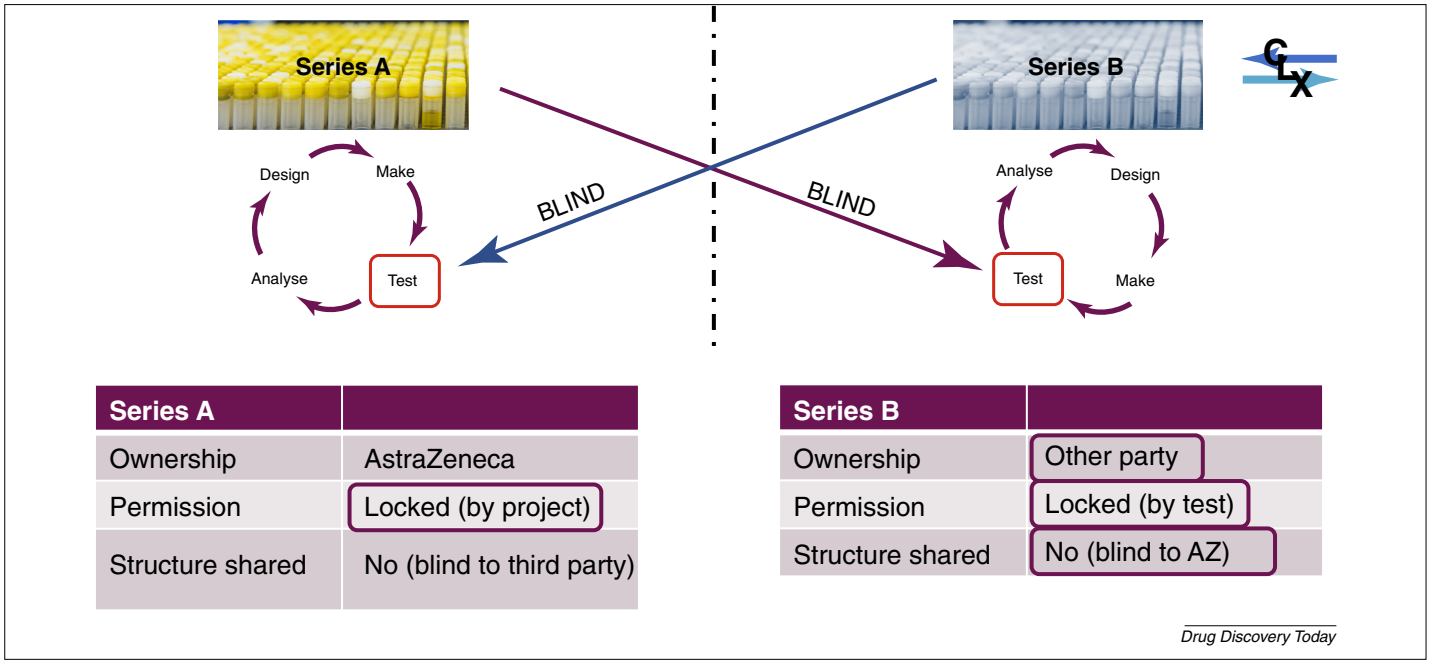
Compound Passport Service (CPS) concepts and definitions. The CPS is based upon assets, which are entities reflecting agreed constraints of compound usage by third parties, and are grouped (together with the rights and rules applying to the assets) in asset rights groups (ARG). For each ARG a number of roles are defined: coordinator – sets up and manages the ARG; delegate – a deputy for the coordinator who can add or remove approvers and add or update assets; approver – a person able to approve or reject requests manually. Within each ARG it is also possible to delineate request rules that define which requests can be automatically approved or rejected. A user in another system that is fully integrated with the CPS becomes a requester when asking for approval to perform a specific action; the CPS responds with the following: 'approved', 'rejected' or 'manual approval needed'.

When placing test requests in a dedicated in-house requesting tool, the CPS is called and the sample status is displayed. Users might be presented with a warning that corresponding samples are subject to approval. Depending on the permission status, the requester can either decide to seek approval or cancel their tentative requests. Extreme cases such as auto-rejection (strictly no testing) or auto-approval (green list) are dealt with instantaneously, whereas manual approvals are immediately sought with daily reminders being sent to the approver until a response is obtained. A feature much appreciated by users is that the system sends an

TABLE 1

Compound Passport Service (CPS) use cases: seven critical scenarios that compound passport needs to be able to service

Use case	Description
Create asset rights group	Set up asset rights group which includes adding compound assets, defining roles and setting up request rules
Update compound assets	Lock or unlock compound assets and modify ownership of compound assets
Request approval	Users in other systems ask for approval to perform an action on a compound (e.g. to perform a specific test). CPS responds with automatic approval, automatic rejection or whether a manual approval is needed
Manual approval	Approver goes into CPS and manually approves or rejects a request
Check compound status	Users in other systems can check status of a compound to learn whether the compound is locked in a collaboration with a partner or not
View asset history	A full audit trail for a specific compound, which collaboration(s) has it been locked in, what requests have been made for this compound, what responses have been issued in CPS, among others.
View request history	View all requests and responses for all assets in a specific assets rights group



Series A	
Ownership	AstraZeneca
Permission	Locked (by project)
Structure shared	No (blind to third party)

Series B	
Ownership	Other party
Permission	Locked (by test)
Structure shared	No (blind to AZ)

FIGURE 2

The Compound Passport Service (CPS) can be used to manage an asset throughout its lifecycle. In this case study, at the start of the collaboration, compounds were tested at AstraZeneca (AZ) and the collaborator structure-blinded. Over the course of research, SAR failed to develop in series A which became of no further interest and was retained by AZ. Series B provided very productive SAR and, through the course of the collaboration, compound properties were updated in the CPS to note initially that the structures had been shared with AZ, then ownership became shared and therefore the series B compounds were unlocked and freely available for test across both organisations (the figure shows the initial collaboration status).

TABLE 2

The three main attributes utilised by the Compound Passport Service (CPS)

Ownership	**Comment**
AstraZeneca	Compound samples belong solely to AstraZeneca
Other party	Compound samples belong solely to third-party partner
Shared	Compound samples are jointly owned by both parties
Permission	**Comment**
Not locked	Compounds can be tested in any assay without restriction
Locked (auto-reject)	Compounds cannot be tested in any assay
Locked (with project green list)	Compounds require manual approval for permission to be tested in any assay except for those belonging to a particular drug discovery project, which are automatically approved
Locked (with named assay green list)	Compounds require manual approval for permission to be tested in any assay except for a green list which is automatically approved, for example important selectivity assays that the partner organisation is aware are being run
Locked (with people green list)	Compounds require manual approval for permission to be tested in any assay except when requested by people on a pre-approved list who understand the collaboration agreement and status of compounds and are trusted to request only permitted tests
Locked (with nonefficacy OK)	Compounds require manual approval for permission to be tested in any assay except nonefficacy tests (physical chemistry assays like solubility, plasma protein binding, metabolism, among others)
Structure shared	**Comment**
Yes	Chemical structures have been seen by both organisations (NB: this is separate to ownership)
No (blind to third party)	Chemical structures have not been seen by collaboration scientists outside AstraZeneca
No (blind to AstraZeneca)	Chemical structures are unknown to AstraZeneca staff

The CPS captures and tracks compound provenance mapped on to three main attributes for each asset: ownership (is it owned by AstraZeneca, the partner or jointly?); permissions (whether the compound is not locked and therefore available to be freely tested, or locked such that automated or manual approval is needed); structure shared (whether or not compound structures are known by both collaborating organisations). The three main attributes can be independently varied for each compound within an asset rights group and status changes can be made (in the form of transitions) over time and as collaborative research develops.

approval email to authorisers giving them the full context of the requests to ensure efficient decision making rather than having to access the CPS to gain the wider context.

Green lists (leading to auto-approvals) are crucial to ensure no impact of sample restrictions on the DMTA cycle. In the case of required manual approvals, the impact of delay is minimised by the creation of project delegates. Green list assays refer to tests agreed with the partner and typically include project assays [primary target, selectivity, *in vitro* drug metabolism pharmacokinetics (DMPK), among others]. Similarly, a green list of requesters typically includes project members who are fully aware of the collaboration. Delegates have the same approval rights as primary authorisers and requests come to all independently, so that the first person to approve them releases the samples for testing.

Interface with other services

The envisioned central role and future extensibility of asset rights management led to the rapid conclusion that the compound passport solution needed to be delivered as a service [12]. The adoption of a service-orientated architecture has provided a flexible and reusable set of business services that provide access to and management of the compound passport database. Additional commodity services provide master data for projects, people and compounds (Fig. 3).

FIGURE 3

Service-orientated architecture provides a flexible and reusable set of business services that provide access to and management of the Compound Passport Database. Additional commodity services provide master data for projects, people and compounds. In this way, the Compound Passport Service (CPS) sits at the centre of a web of independent services controlling compound registration, distribution and test approval. Utilisation of reusable services allows integration with new systems as they become available in the future.

In addition to the above-described interfaces to test request management, the CPS has the potential to extend into control of compound distribution, through linkage to the services managing compound provisioning and ordering. The information in the CPS in many cases supersedes the more general legacy compound restriction information and positions the CPS as the future replacement for this component of the current compound registration platform. This will benefit the Compound Management Group which is responsible for sample shipping and receipt associated with all open innovation collaborations. The Approval Request Service within the CPS will provide a superior mechanism to initiate, track and record the approval decisions required for any such compound sharing. The ability to call the CPS for ARG assignment will allow newly synthesised compounds to be classified at the point of registration, based on criteria such as project or synthesising organisation, avoiding the need for retrospective assignment.

The information that the CPS holds on the asset rights status for compounds and samples can be exposed in downstream query applications, providing scientists with up-to-date information upon which to base DMTA decisions. The CPS is accessed by single sign-on, federated security and has the potential for externalisation at a future date to allow alliance partners visibility of their assets history.

Impact on open innovation

Since the CPS was launched in 2014, it has been uploaded with metadata relating to 15,000–20,000 assets, and compounds are being added almost daily. The ability to track ownership rights along with more-efficient test approval has already enabled faster and more-efficient approval flows. Additionally, in considering whether to unblind the structure of a HTS hit series, we have also been able to identify that more than one external party had an interest in the chemical equity. Based upon an understanding that SAR would probably diverge as potency against the different targets was optimised, we have been able to adopt a risk-management approach to allow both partners to proceed with the investigation and possible optimisation of the shared chemical start-point.

AZ is a leading participant in a number of open innovation initiatives: IMI European Lead Factory [13]; CRT Metabolism Alliance [14]; MRCT Screening Alliance [15]; and the Open Innovation Portal [16] – the CPS underpins all these efforts as well as several additional collaborations. It is widely forecast that the trend towards open innovation and compound sharing collaborations will accelerate [6,17], and therefore the ability to track multiple-owner compounds will increase in importance. Anecdotally, the absence of an open-source or vendor package to facilitate the sort of tracking enabled by the CPS actually deters some university academic groups from making research materials available to collaborators. Our ability to keep track of how we 'keep our promises' with partners as well as the capacity to record dates of structure sharing accurately is seen as attractive by potential new partners. This in turn is opening the way to help us subtly tune collaborative research plans to suit the needs of the partner and the work plan envisaged.

As already noted, it is now much easier for chemists to track the status of HTS work-up activities, in particular the recording

of sample dispatch and structure sharing with partners increases the confidence of scientists working in complex partnerships. Project scientists value being able to update sample metadata dynamically to show that compound structures have now been shared, while maintaining the record of ownership and testing rights.

The service maintains an exhaustive audit trail of all compound metadata changes, approval and/or rejection tracking and thus provides an element of compliance monitoring; in consequence we have been able to provide partners with whom we hold large collections with a comprehensive record of requests for 'their' samples.

Concluding remarks and future outlook

The flexible service-orientated architecture underlying the CPS makes the writers' task of predicting the full extent of its future utility difficult. However we already foresee some near-term developments and, because the CPS has interfaces designed to enterprise standards, this will allow future requesting systems to interface to the CPS solution when it suits them in their own system lifecycles. The success of cloud-based solutions that have securely placed compound proprietary information outside company firewalls for workflow tracking; for example chemTraX© [9], for virtual screening [18] and open source- and collaborative drug discovery (OSDD and CDD) [19,20], is an interesting development that is conceptually a good fit for the CPS for externalisation at a future date to allow alliance partners visibility of their asset transactions. The emergence of open innovation portals could in turn create new business models for compound sharing that might in turn drive the development of the CPS in new directions. An interesting future possibility is that the CPS can in fact run completely structure-blind and so it need not be confined to small-molecule assets. The CPS could be extensible to almost any intellectual or proprietary asset that is shared with collaborators in an open innovation context. Oligonucleotides, proteins, molecular biology tools, cells, tissues, transgenics, even software could be managed using the service, albeit that the CPS might soon seem too narrow a term to use.

A final exciting trend is the emergence of cross-industry, national and international compound collections. Database service solutions will evolve around these national assets and these could also develop in ways that aim to maximise researcher impact through increased collaboration, technology transfer and commercialisation while simultaneously lowering barriers to collaboration and licensing through reduction in the administrative thicket of patents, confidential disclosure agreement (sometimes termed non-disclosure agreement (NDA)) (CDAs) and material transfer agreement (MTAs) [21,22]. It is conceivable that flexible service-orientated solutions such as the CPS will be able to adapt quickly to make direct links to national compound databases, allowing collaboration partners to view the status of their compounds in real-time within what appears to the user to be a single system.

As has been noted elsewhere [6,23], we are moving towards a more open world, in which organisations need to collaborate to thrive. The old, linear paradigm where each player's position was clearly defined has evolved into a dynamic network of nontraditional partnerships in which compounds, data, expertise and knowledge are shared. Our development of the CPS is one mechanism that provides support and clarity where the roles of industry, academia, charities and research funders in innovation are increasingly overlapping.

Acknowledgements

The authors are very grateful to Olga Seliverstova and her team at EPAM Systems in St Petersburg for their contributions to this work.

References

1 Paul, S.M. *et al.* (2010) How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nat. Rev. Drug Discov.* 9, 203–221
2 Ganellin, C. (1995) Historical approaches to drug discovery. *Ars Pharm.* 36, 329–335
3 LaMattina, J.L. (2011) The impact of mergers on pharmaceutical R&D. *Nat. Rev. Drug Discov.* 10, 559–560
4 Jorgensen, W.L. (2012) Challenges for academic drug discovery. *Angew. Chem. Int. Ed.* 51, 11680–11684
5 Chesbrough, H.W., ed. (2003) *Open Innovation: The New Imperative for Creating and Profiting from Technology*, Harvard Business School Press
6 Pigott, R. *et al.* (2014) *Shaping the Future of Open Innovation: A Practical Guide for Life Sciences Organisations*. Available at: http://www.wellcome.ac.uk/About-us/Publications/Reports/Biomedical-science/ (accessed 28.03,15)
7 Krämer, K. (2015) Crossing the valley of death. *Chem. World* 12, 49–51
8 Kogej, T. *et al.* (2013) Big pharma screening collections: more of the same or unique libraries? The AstraZeneca–Bayer Pharma AG case. *Drug Discov. Today* 18, 1014–1024
9 Andrews, D.M. *et al.* (2014) Collaborative practices for medicinal chemistry research across the big pharma and not-for-profit interface. *Drug Discov. Today* 19, 496–501
10 Jordan, A.M. *et al.* (2015) Rethinking 'academic' drug discovery: the Manchester Institute perspective. *Drug Discov. Today* 20, 525–535
11 Plowright, A.T. *et al.* (2012) Hypothesis driven drug design: improving quality and effectiveness of the design-make-test-analyse cycle. *Drug Discov. Today* 17, 56–62
12 Service oriented architecture: what is SOA? In: *The SOA Source Book*. Available at: http://www.opengroup.org/soa/source-book/soa/soa.htm (accessed 28.03.15)
13 Mullard, A. (2013) European lead factory opens for business. *Nat. Rev. Drug Discov.* 12, 173–175
14 Jones, N.P. and Schulze, A. (2012) Targeting cancer metabolism – aiming at a tumour's sweet-spot. *Drug Discov. Today* 17, 232–241
15 Osherovich, L. (2012) AstraZeneca's compound interest. *SciBX* 5 http://dx.doi.org/10.1038/scibx.2012.2
16 Available at: http://openinnovation.astrazeneca.com/ (accessed 28.03.15)
17 Melese, T. *et al.* (2009) Open innovation networks between academia and industry: an imperative for breakthrough therapies. *Nat. Med.* 15, 502–507
18 Available at: http://www.enterprisetech.com/2014/07/10/aws-talks-enterprise-wins-adds-zocalo-collaboration/ (accessed 28.03.15)
19 Bhardwaj, A. *et al.* (2011) Collaborative tools to accelerate neglected disease research: open-source drug discovery model. In *Collaborative Computational Technologies for Biomedical Research* (Ekins, S., ed.), pp. 321–334, John Wiley & Sons
20 Ekins, S. *et al.* (2011) Pioneering use of the cloud for development of collaborative drug discovery (CDD) database. In *Collaborative Computational Technologies for Biomedical Research* (Ekins, S., ed.), pp. 335–361, John Wiley & Sons
21 Hall, B.H. *et al.* (2013) *A Study of Patent Thickets*. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/311234/ipresearch-thickets.pdf (accessed 28.03.15)
22 Wesley Hill, W. (2014) Open innovation's answer to the challenges of patent overpropertization. *J. Intellect. Prop. Law Pract.* 9, 730
23 Papadatos, G. and Overington, J.P. (2014) The ChEMBL database: a taster for medicinal chemists. *Future Med. Chem.* 6, 361–364