As part of the Digital Single Market strategy, the European Open Science Cloud (EOSC) can support the progress of science and the maximisation of data discoverability, accessibility, preservation and reuse, if it succeeds in achieving a seamless integration of the many and varied distributed resources researchers are able to use in the pursuit of their research—in particular data storage, compute cycles, and the packaged combinations of data and compute we term services—in a way that is easy to use.

Many of the resources and services needed for the EOSC already exist; while technical challenges remain, the main barriers are ones of policy and concern funding, lack of interoperability, access policies and coordinated provisioning.

EUDAT sees the following five challenges as the most pressing for the success of the EOSC. Some of these map well to the challenges already identified by the high-level expert group (HLEG); some are new.

**Challenge 1** (cf. HLEG A1): Lack of consensus on tools and APIs, and lack of authoritative persistent identifier schemes.

This principal technical challenge to the free flow of data arises from too many solutions rather than too few. EUDAT believes that progress hinges on agreements between significant players on solutions which are “90% good enough”, with implementation built upon a willingness to compromise. EUDAT’s collaborative data infrastructure (CDI) strategy is defining common standards and policies for data access and ingest APIs (based on industry standards); for persistent identifier behaviour (with the Handle system and EPIC API); and common baseline metadata for all identified objects (following the OpenAIRE guidelines). These will be adopted across the consortium, with commitments cemented in a formal partnership agreement in the first half of 2016. We believe this will contribute significant momentum towards well-defined, stable APIs for the EOSC.

**Challenge 2:** Lack of consensus on authentication, authorisation, single sign on and common user identity.

The storage and computing resources that underpin the EOSC will not be owned by the EOSC. They will be owned by national or regional governments, or by private service providers (cf. Challenge 4). Decisions about which users are authorised to access these resources will thus be localised, and to support ease-of-use across multiple resources in multiple administrative domains, common agreement on, first, user identity and, second, accepted authentication mechanisms is essential. The solutions here will come from agreement between European (and global) identity providers through initiatives like eduGAIN. EUDAT is lending momentum to European single sign-on through adoption of emerging technologies like the Unity Identity Management system.

**Challenge 3:** Sustainability of data-centric research infrastructure.

Perhaps the biggest policy challenge for the long-term (or even medium-term) preservation of digital research data is sustainability. In providing perishable goods like computing cycles, private service providers have a key role in offering flexible cloud-based solutions; the same is much less true for the long-term storage of expensive, or potentially irreplaceable, research data. Private storage providers
may—indeed should—have a role, but only as part of a wider public strategy. Especially in the era of open data, where data generated by public funds should be publicly available (at little or no cost to the end-user), scientific data infrastructure is either publicly funded or not at all.

EUDAT believes that the European Union has a pivotal role to play here in funding infrastructure that is persistent, multi-disciplinary and trans-national (cf. also Challenge 4). Only a publicly funded and publicly governed EOSC can guarantee persistence and sustainability, and ensure that outcomes are driven by scientific excellence and societal needs.

Challenge 4 (cf. HLEG B1-4 and B8): Fragmentation, across discipline, national and regional boundaries, of infrastructure funding and business models.

A natural consequence of funding excellence in science is deep specialisation; while appropriate for basic research this approach is far less suited to general-purpose compute or data infrastructure, either in terms of funding or of career development for practitioners. A similar argument can be made with respect to national or regional funding. As noted under Challenge 2, research infrastructure is typically procured nationally against national priorities and with varying degrees of flexibility in its use for trans-national purposes.

This challenge should be accepted as a constraint rather than seen as a problem to be solved, at least in the medium term. The EOSC must be developed in a way that accepts and works with this fragmentary landscape, recognising that underlying resources are distributed and have very different ownership and business arrangements. EUDAT has adopted a similar flexible, pragmatic approach to its own business modelling, planning in terms of a “regulated market” rather than a purely free market or, worse, a monopoly supplier. Market forces can drive innovation, but continuous innovation can result in continuous disruption and instability. By focusing strongly on interoperability “in the middle” the EOSC will create the necessary fixed point of stable standards and APIs, while accommodating flexible resource provision beneath, and promoting scientific and technological innovation above.

Challenge 5 (cf. HLEG B7): Europe-only vs. global approach.

The EOSC must be a European implementation of a global OSC model rather than a Europe-only construct. That global model for Open Science infrastructure must recognise—indeed it must welcome—participation by global private service providers, but its development must be rooted in the needs of research. The Research Data Alliance must take a leading role in fostering development of the OSC model, and Europe’s continuing strong leadership within RDA provides excellent opportunities to set the agenda. EUDAT's close links with RDA and our role as a natural implementer of RDA recommendations will add significant weight to European leadership in Open Science infrastructure.