SAUCE: Asset Libraries of the Future

Jonas Trottnow
Filmakademie
jonas.trottnow@filmakademie.de

William Greenly
DNEG
wmg@dneg.com

Christian Shaw
DNEG
cshaw@dneg.com

Sam Hudson
Foundry
sam.hudson@foundry.com

Volker Helzle
Filmakademie
volker.helzle@filmakademie.de

Henry Vera
DNEG
hyv@dneg.com

Dan Ring
Foundry
dan.ring@foundry.com

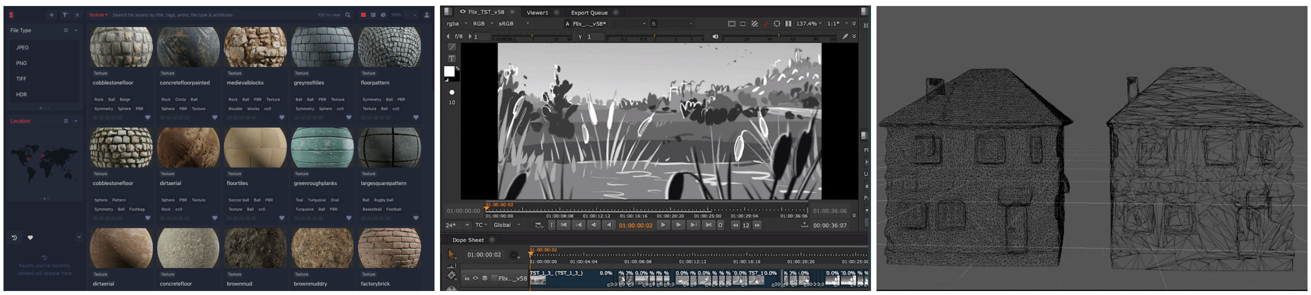

Figure 1: Components and use cases of the SAUCE asset library system. Left, an example search for materials in DNEG's asset search & retrieval tool. Centre, building a timeline in Foundry's "Hiero" tool by retrieving panels from asset storage system. Right, example transformation of stored assets for use in Filmakademie's virtual production system, in this case mesh complexity reduction.

ABSTRACT

Storage and retrieval of production assets is vital for every modern VFX and animation facility. From the volume of assets being stored to the constantly changing variety and richness of the asset data, efficiently storing, indexing, finding and retrieving the assets you want is a growing challenge. This paper discusses some of the requirements of modern asset storage systems for VFX and animation, introducing two systems that were built to address these challenges as part of the collaborative EU funded "SAUCE" project; DNEG's search and retrieval framework, and Foundry's back-end asset storage. It also presents example use cases of the asset library from Filmakademie's experiments in virtual production, demonstrating more artist focused and task centered systems that enable greater asset re-use.

CCS CONCEPTS

• **Information systems → Information retrieval**; **Multimedia databases**; *Record storage systems*; *Data encoding and canonicalization.*

KEYWORDS

asset search, asset storage, asset library, asset management, asset retrieval, information retrieval

ACM Reference Format:
Jonas Trottnow, William Greenly, Christian Shaw, Sam Hudson, Volker Helzle, Henry Vera, and Dan Ring. 2020. SAUCE: Asset Libraries of the Future. In *The Digital Production Symposium (DigiPro '20), July 18, 2020, Virtual Event, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3403736.3403941

1 INTRODUCTION

Studios collect and generate large volumes of data and require that data be stored and retrieved efficiently. At its core, that is the minimum that an Asset Library must do.

Today, the demands on these systems are greater, not only due to the scale of the data, but the ways in which the data is managed, replicated, indexed and delivered. We work in a visually and semantically rich industry where the quality and format of images are as important as their content. VFX and animation are industries of constant technological evolution where new formats and standards replace old ones, and yet old ones never fully go away. They are also industries where intelligent reuse of previous assets can improve efficiency and lower cost.

The work presented here was developed as part of "SAUCE", a collaborative EU Research and Innovation project between Universitat Pompeu Fabra, Foundry, DNEG, Brno University of Technology, Filmakademie Baden-Württembergs Animationsinstitut, Saarland University, Trinity College Dublin, and Disney Research to create a step-change in allowing creative industry companies to re-use existing digital assets for future productions [SAUCE 2020]. The goal

DigiPro '20, July 18, 2020, Virtual Event, USA

ACM ISBN 978-1-4503-8034-8/20/07…$15.00
https://doi.org/10.1145/3403736.3403941

of SAUCE is to produce, pilot and demonstrate a set of professional tools and techniques that reduces the production costs of enhanced digital content. In particular, targeting the creative industries by increasing the potential for re-purposing and re-use of content as well as providing significantly improved technologies for digital content production and management.

This paper is split into the three core technology areas of modern asset libraries, each explored by a different partner. In Section 2, Foundry's work on asset storage and access is introduced. Section 3 presents DNEG's framework for intelligent indexing, search, transformation and retrieval of assets. In Section 4, the argument is made for asset transformation through an example virtual production. Finally, all of the systems are brought together and discussed in Section 5.

2 FOUNDRY: STORAGE & ACCESS

The first stage of any asset library system is deciding how and where to store the data, as well as providing controlled access to it. There are many ways of doing this from simple file-system based network access storage to relational object database storage systems. From here, the requirements of a VFX or animation studio will start to define the storage system. In this section, the design decisions and features of the developed storage system are presented, all of which have come from customer requirements.

Core Concepts. There are some common core concepts around assets which most asset libraries must adopt. The first is that an asset relates to a single uniquely addressable entity. An asset must also be version-able, where a counter for that asset is incremented every time a new version of the asset is published or checked-in. This also means that when multiple artists are working on the same asset concurrently, users must be made aware this is happening, and there needs to be some form of conflict resolution.

Spirit of Assets. One of the most interesting lessons learned from customers is that creative people don't want to work with files. Once imported into the storage system the user should not need to know about what the file is or where it's stored.

Seamless Format Transform. Artists should also not need to know what data type or format the asset is stored as, only what they can retrieve the asset as. This ensures both forwards and backwards compatibility with various file formats and is enabled by a "Transform Engine". This is one of the key features of this work. The Transform Engine implements a plug-in architecture to enable parsing and transformation of assets. For example, if an asset was imported as an FBX file, it could be retrieved as an OBJ for compatibility with legacy software. Similarly, an Alembic file could be automatically converted and conformed to the studio pipeline's USD format. This will be covered in more detail later in Section 4

Arbitrary Data Types. The variety of file types in VFX and animation studios is staggering. A key requirement of an asset library is that imported assets can not be limited to known file types. Each asset should be stored regardless of file type, with the provision that if a file-type is known to the system it may be operated on.

Accessibility. When an asset is requested from storage, it needs to be easily accessible. This has been implemented through a simple and well-documented REST API. Modern pipelines and infrastructure can be heavily customised therefore having a language- or software-agnostic way to retrieve assets is vital.

Resilience & Performance. The storage system must provide redundancy and load-balancing to guarantee performance. This has been achieved through server-side replication policies. It was found to be important to offer a scale of redundancy. For example, a "copy-all" policy ensures that data is copied to all site, ensuring high performing multi-site operation and high data redundancy. A "don't copy" policy ensures data is not automatically moved between servers unless requested by a client, helping to reduce unnecessary replication of assets. This is very useful when multiple sites (or even multiple servers within a site) are being used for different shows for example. To track the assets and store their metadata, the system relies on a single database being visible to all servers. This is sufficient as in general the ratio of publish vs. check-out (write vs. read) operations is very low.

On-prem & Cloud. The cloud is becoming more and more attractive for collaborative data access, particularly where data needs availability, redundancy and performance guarantees. While concerns around cloud security are easing, there are still cost, knowledge, ownership and integration barriers that mean customers often prefer to keep data locally. The presented system has been developed to operate both locally on a facility's existing infrastructure as well as in the cloud, or even a hybrid of the two.

Turn-key. The storage system needs to work in modern VFX and animation studio pipelines. This means all of the above features must be delivered through the storage system alone. For example, while cloud-based load-balancing or replication services could be enabled by running the storage system in the cloud, some functionality of these features must also exist without third-party services. The idea is that the system should be able to work independently, and work easily for both large facilities as well as small studios with no IT departments. This has been achieved largely through containerisation and understanding customer infrastructure to minimise deployment issues.

To visualise how all of the above components and features fit together, a summary of the asset storage system is shown in Figure 2. With the asset storage defined, the system for indexing assets for search and retrieval is introduced.

3 DNEG: SEARCH & RETRIEVAL

The problem of how to find the data you want is universal. Traditional information retrieval systems operate on text, where indexing of text and language structures are well understood. However the demands of the VFX and animation industries are greater. The visual nature means that these systems need to move beyond text. The volume of data produced means they need to move beyond hand-labeling and curation of metadata. And importantly the cost and value of the assets created is so high they must be easily and reliably retrievable.

A summary of some of the challenges facing DNEG and its existing systems were as follows:

- How to curate large archives of data without employing significant costs in human resources?

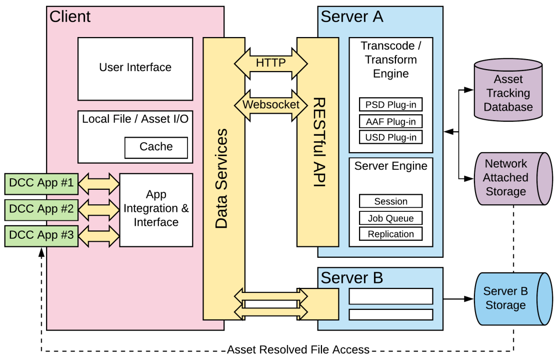

Figure 2: Diagram of the developed back-end asset storage system, and its relationship to the client. Each server maintains a connection to a shared database, which can be replicated to provide redundancy and avoid being a single point of failure. Each server can also maintain its own storage or access shared networked storage. The system can be integrated into DCC apps by interfacing with the local client.

- How to curate things consistently and uniformly to enable accurate search results?
- How to support the classification and enrichment of data spanning a variety of intersecting and ever expanding domains? In the VFX world, new types of assets are continually being created and any system needs to support the resultant domains without substantial re-architecture.
- How to enable domain experts to contribute vocabularies to support the enrichment and classification of those domains?
- How to provide a rich search experience that can support interrogation of a plethora of asset types in a consistent and intuitive way?

The SAUCE project seeks to address some of these issues by providing an extensible and scalable framework to enable users or third-parties to contribute vocabularies, classifiers and transformation services that provide and enrich asset metadata, along with a rich search and user experience for retrieving these assets. The main building blocks are described below.

3.1 Published Vocabularies and Descriptors

One of the key building blocks is a framework to enable users or third-parties to contribute vocabularies that describe key concepts and terms that are interesting to artists, for a particular type of asset or production scenario. Technically speaking, the framework leverages a variety of Linked Data Standards, namely Resource Description Framework [Candan et al. 2001; RDF 2014] and Web Ontology Language [OWL 2012], in order to canonicalise asset descriptors and semantics. By canonicalising the way vocabularies are represented, the framework allows contributors to align vocabularies and concepts with other vocabularies or concepts, in a decentralised fashion. This provides a scalable and extensible means for contributors to contribute autonomously, without the constraint of centralised consensus currently prevalent in existing search and asset management systems. These published vocabularies also form the basis of search and filter, and are the foundations of the interoperable and composable classification and transformation framework detailed in the next section.

In addition to vocabularies provided by partners, the framework also leverages open vocabularies, schemas, knowledge bases and dictionaries, including [Schema.org 2012], [WikiData 2012] and Wordnet [Miller 1995] to name a few, with the support to extend dictionaries to include industry specific terms and nomenclature.

3.2 A Scalable and Extensible Classification and Transformation Framework

To reduce the effort that would normally be required to curate and classify large quantities of asset data, the framework provides a means for users and third-parties to contribute classifiers and enrichers [Greenly 2018]. These classifiers have been trained to classify and enrich different types of assets, extracting semantic and intrinsic metadata from assets. The majority of these classifiers have been pre-trained using machine learning and neural networks. One of the key goals of the framework is to enable interoperabilty and chaining of classifiers, thereby enhancing and augmenting their capabilities and maximising the value of any individual contribution.

The framework also provides support for transformation capabilities [Greenly 2019] to be advertised and integrated into both search and classification. This allows more options for classification and also expands search results to include assets that are relevant after a transformation. This greatly increases the chances of asset reuse, generating more value from existing assets.

Another important feature is the ability to chain together classifiers and transformations, both implicitly and explicitly. By declaring individual capabilities, the framework can implicitly work out achievable states of assets from a series of transformations or classifications, without advertence or intent; whilst explicit chaining provides a composable means to curry transformations and achieve desired states through optimised routes.

On a more technical level, the framework heavily leverages containerisation. Similar to the asset storage system, this provides a high degree of abstraction, modularity, dependency and run-time isolation, which helps the system's scalability and re-usability. In particular this enables classifiers and transformation components to be deployed and run in a variety of environments with different constraints, configurations and profiles. This was especially important within DNEG's infrastructure, which has strict security constraints, combined with a hybrid on-premise/cloud footprint. By containerising the components, it provides a consistent architecture for building and testing components, whilst being able to deploy and run them in different target environments, with different constraints.

3.3 An Intuitive and Faceted Search

To support the efficient retrieval of assets for reuse, a new search user interface and search API is being developed. This interface provides faceted search, which allowed artists to find different types of assets in different ways depending on personal preference and production context. The visual design patterns and user experience

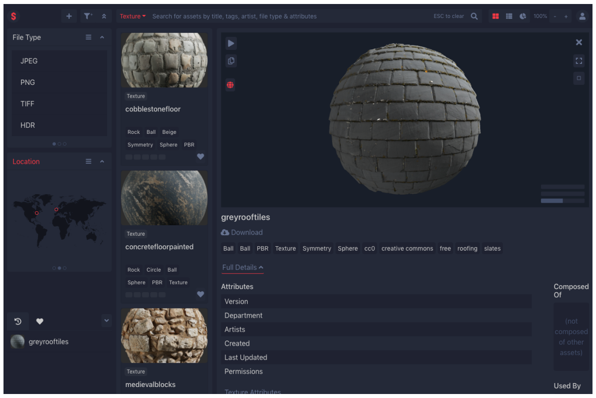

Figure 3: Detail view of an example search for materials in DNEG's system. Materials are rendered by system as images for artist-friendly visualisation, and the various hand-labelled or automatically classified tags are shown. Also shown are the various filters, such as region or file-type.

architecture underpinning this interface has to be extensible to support a variety of asset types, filters and navigation components. To this effect a user interface vocabulary is being developed that aligns asset descriptors and concepts to user interface components, thus allowing for a dynamic, extensible and fluid user interface. Furthermore, by faceting the search, artists are able to visualise and filter results in different ways for different types of asset or production context. An example of full and detailed search for materials can be found in Figures 1 and 3.

The user experience was developed using standardised web technology, which allowed component widgets to be embedded and reused in a variety of legacy in house applications, thus augmenting and enhancing existing applications and experiences and enabling a seamless experience for artists.

4 FILMAKADEMIE: TRANSFORMATIONS & USE CASES

The context in which an asset is created or consumed can vary greatly. As example, an asset modelled and rigged for a VFX sequence, may need to be reduced in complexity or have its textures transformed for use in a virtual production scenario.

An often underestimated overhead for virtual productions is this asset preparation. Assets generally need to be available in high detail, to be able to judge visual quality, and in a reduced detail version used for real-time purposes. This is the workflow for productions involving “VPET” - Virtual Production Editing Tools - developed by Filmakademie Baden-Württemberg (FA), shown in Fig 4. As part of SAUCE, the asset pipeline and workflows have been targeted for virtual production and use with VPET.

In a virtual production context, descriptive metadata greatly simplifies the setup and authoring of an initial scene. For example, assets that are explicitly suited to virtual production can be found conveniently, such as real-time capable rigs, meshes with fewer polygons etc. Metadata also serves as a guide for the transformation part of the framework: if an asset does not fulfill all requirements

Figure 4: The VPET framework can retrieve its real-time optimised scene from any host application, or ingest USD files containing various LOD's generated by the system.

of a search, the framework can look up if a transformation plug-in exists that is capable of automatically transforming the asset to fulfill all search criteria. Based on the labels, a scene can be transformed for use in virtual production environments, e.g. in VPET. Through the labels, movable objects like cars, benches etc. can be separated from static parts like houses. These movable objects are then marked as editable in the VPET system, enabling a user to select and modify it in real-time on set with the VPET tablet clients. In addition, the labels can be used to provide scene understanding for example to the animation solving engine in order to generate scene aware animations.

Since the VPET tablet client is used as a remote control for the 3D scene, visual quality on these mobile devices is not the key target. VPET allows the usage of different versions of the same asset on the client and the host. This means that a user can interact with the real-time asset on the tablet, but all updates are applied to the high quality version on the scene host, which thereby is capable of providing a high quality preview of the final look.

This approach requires two versions of the same asset. Creating them manually is time consuming. Numerous algorithms are able to reduce the complexity of an asset automatically while maintaining shape and textures. The search and transformation framework provides the possibility to integrate such a transformation as plug-in, thereby simplifying the preparation of a virtual production and making it more attractive for smaller studios or individuals. In addition the framework can transform asset formats. For example, a USD scene of several assets at various LODs can be automatically generated, composed and delivered. It could also provide metadata, containing domain and scene specific information or variations of assets. Or more simply, the transformation plug-in could provide a translation between the structure of one pipeline and another, something which is very important when migrating assets from a film or episodic pipeline to a virtual production one. More and more DCC tools are supporting USD, including VPET. File transformations such as these simplify the pipeline in virtual production scenarios.

A concrete application is an upcoming virtual production based VFX car shoot realised by FA, where a real car will be placed in front of a large, high-resolution LED wall. The large screen and additional

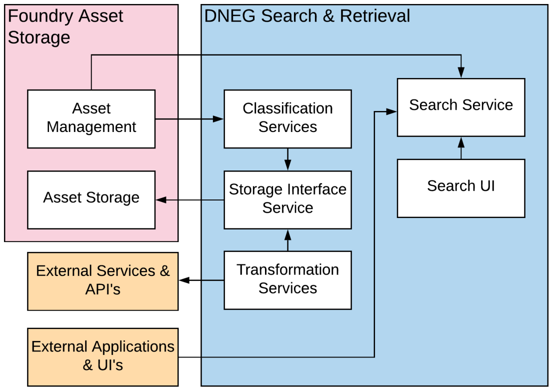

Figure 5: SAUCE Asset Library Architecture, illustrating how DNEG's asset search and retrieval system integrates with Foundry's asset storage system.

smaller panels serve as backdrop and lighting. Through the search and transformation framework the imagery being displayed on the wall can be quickly exchanged on set. Director wishes can be quickly addressed, sourced and transformed from existing asset databases. 3D elements (also rendered on the LED wall) can be placed in the middle ground between car and background. Camera effects and compositing will be handled on set in camera. The aim is to use VPET to enable users to interact and modify the 3D elements provided by the search and transformation framework in real-time.

5 SAUCE ASSET LIBRARY & CONCLUSION

This paper presented a pair of systems for asset storage, and asset search and retrieval. Together they comprise the SAUCE Asset Library system. To understand how they relate to each other, an architecture diagram is shown in Fig 5.

Looking across the asset indexing landscape, the two closest systems are the impressive works of [Zorroa 2020] and [Yadle 2020]. Avoiding a direct feature comparison, the key differences in approach between those systems and that of SAUCE is that 1.) SAUCE focuses exclusively on VFX and animation, and 2.) SAUCE strongly encourages customisation through user plug-ins. While the commercial offerings also offer machine learning in their tools as a way to enrich data, a lot of what is needed in VFX and animation is often simpler. For example, a tool to pull proprietary lens metadata from a camera file, or a script to say how different a pair of Nuke scripts are. For these scenarios a more domain specific yet flexible and customisable system is preferred.

A key contribution of this paper is the detailing of assert storage system requirements targeted for VFX and animation and derived from real customer feedback. Although the presented system is implemented in Foundry's "Flix" story development tool, the requirements serve as a checklist for any facility wanting to develop their own asset store, using the implementation shown in Fig 2 as a starting point.

The second contribution of this paper is the presentation of DNEG's search and retrieval system, answering industry specific problem statements. In particular it demonstrates how to reduce the cost of human interaction, and the value in supporting system interrogation by a plethora of data types. It also illustrates how to make the most of rich data and enrich "poor" data through an extensible framework for classification and enrichment.

From previous experience [Ring et al. 2019] it has been found that VFX and animation studios are not only comfortable with using machine learning, but have an appetite for actively developing their own models and tools. Having an extensible system that can provide tailored classification to the particular needs of the studio is vital. For example, being able to detect colour checkers or slates [Kleiman et al. 2019].

Lastly, Filmakademie's virtual production experiments demonstrate the value of the transformation frameworks. Virtual production is becoming increasingly popular, and the richness and variety of the data that is both generated and consumed means there needs to be a way to simplify the transformation of assets for easy re-use in different contexts. This is vital in making virtual production more accessible for smaller studios or individuals.

As mentioned in Section 4, for future work the above systems will be tested in an upcoming "LED wall" based virtual production. The shoot will test whether assets in storage can be efficiently searched, retrieved and transformed on-set to match the creative needs of the director, and the technical requirements of LED walls.

ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No 780470.

REFERENCES

K. Selçuk Candan, Huan Liu, and Reshma Suvarna. 2001. Resource Description Framework: Metadata and Its Applications. *SIGKDD Explor. Newsl.* 3, 1 (July 2001), 6–19. https://doi.org/10.1145/507533.507536

William Greenly. 2018. Smart Search Framework POC. https://www.sauceproject.eu/dyn/1549296689552/D4.1.pdf.

William Greenly. 2019. Basic framework to enable asset transform. https://www.sauceproject.eu/dyn/1568799474026/D5.3.pdf.

Yanir Kleiman, Simon Pabst, and Patrick Nagle. 2019. Boosting VFX Production with Deep Learning. In *ACM SIGGRAPH 2019 Talks* (Los Angeles, California) *(SIGGRAPH '19)*. Association for Computing Machinery, New York, NY, USA, Article 24, 2 pages. https://doi.org/10.1145/3306307.3328208

George A. Miller. 1995. WordNet: A Lexical Database for English. *Commun. ACM* 38, 11 (Nov. 1995), 39–41. https://doi.org/10.1145/219717.219748

OWL. 2012. OWL 2 Web Ontology Language. https://www.w3.org/TR/owl2-overview/. [Online; accessed 07-May-2020].

RDF. 2014. RDF 1.1 Primer. https://www.w3.org/TR/2014/NOTE-rdf11-primer-20140624/. [Online; accessed 07-May-2020].

Dan Ring, Johanna Barbier, Guillaume Gales, Ben Kent, and Sebastian Lutz. 2019. Jumping in at the Deep End: How to Experiment with Machine Learning in Post-Production Software. In *Proceedings of the 2019 Digital Production Symposium* (Los Angeles, California) *(DigiPro '19)*. Association for Computing Machinery, New York, NY, USA, Article 6, 5 pages. https://doi.org/10.1145/3329715.3338880

SAUCE. 2020. Sauce: Smart Assets for re-Use in Creative Environments. https://www.sauceproject.eu/. [Online; accessed 29-April-2020].

Schema.org. 2012. Schema.org: A collaborative community creating, maintaining, and promoting schemas for structured data on the Internet. https://schema.org/. [Online; accessed 07-May-2020].

WikiData. 2012. Wikidata: A free and open knowledge base that can be read and edited by both humans and machines. https://www.wikidata.org/wiki/Wikidata:Main_Page. [Online; accessed 07-May-2020].

Yadle. 2020. Yadle: Intelligence Improves Business Results. https://www.yadle.com/. [Online; accessed 29-April-2020].

Zorroa. 2020. Zorroa: Intuitive automated machine learning for your business. https://zorroa.com/. [Online; accessed 29-April-2020].