

- SL: 00:28 Oxdb is an online movie database. It's an online repository of cinema. So the objects in it are films spanning a hundred years, more than a hundred years of the history of cinema. International, independent off-off cinema. Also quite a few objects that are not really part of the usual canon of cinema. So that's the kind of content and the idea is that all of these movies, these films exist as films in the database but there is a lot of information associated with them. Not just in the sense of traditional metadata for the films, which it all includes, but there is time-based annotation with them. And since one of our main ideas when we approached this is that we thought that in many online repositories for film was never properly treated as a digital object.
- SL: 01:32 We thought it was evident to develop new ways of looking at film online as a digital database. They are different from just looking at an item. We're looking at an item plus data, you know. The core kind of formal feature of it that are a number of ways of visualizing the contents of the film, and seeing a lot more than we could usually see at one glance, ways of searching through film that brings it directly into the film directly onto a frame or the sequence of frames. Ways of recombining films, ways watching results as movie again or as a kind of an assemblage of clips. So while it is at its core a kind of classical database or repository of film, it was always intended to confuse, or question, the way we deal with an archive.
- JG: 02:31 In some ways it's also not so much a database of films that leads to the object, the film. The films really become part of the database and through forms of recombining, searching and working with the material, the films become in some ways a material that you work with as you navigate the database.
- SL: 02:52 We didn't try to foreground the single item, but that we try to make visible is context, links and we tried to dissolve the single item, the single work a bit and find ways of bringing things together that are usually not seen together so easily. But at the same time, Oxdb depends a lot on its own context, which is space in terms of people with whom we are working on very

similar, related sister projects. But that it's also context in terms of what was the place we developed it in, who were the people who surrounded us, why we were doing it and what kind of actual physical spaces in Berlin, in Bombay and also budgets leak into.

- SL: 03:50 So we started developing Oxdb at Bootlab in Berlin in 2007. Bootlab was a kind of autonomous media lab in the broadest sense. It hosted a lot of initiatives in and around arts, technology, politics in a kind of Berlin fashion. Its traces go back to Internationale Stadt and Botschaft eV.. Mikro eV. Was kind of instrumental in establishing it. It opened in 2000. That's where we met also. We were both members very early on and it was a place where people worked on projects. It changed a bit during its course in a sense that some people have brought in funding and funded project. It didn't have funding on its own, but there were projects in it that were funded. And when we started Oxdb that was kind of formally part of the project we were running in 2007 called the "Oil of the 21st Century", a somewhat meandering, wide-ranging project about issues of intellectual property. We had a vague idea in the beginning that we didn't just want to do talks and conferences and traveling and writing, but we also wanted to produce something concrete. So we had some time and space to produce an archive.
- JG: 05:23 We started with a more general database of digital objects and wanted to find some way of representing different media types and different things. We then realized that it was very hard to really work with the material in a way that allows for interesting forms of interacting, for annotation, but also to show them together and to leave the individual objects behind it and create something that brings the different things together. And so then we started to reduce the scope to a film database as opposed to one about digital objects that could range from images to music to newspapers, to books, to films.
- SL: 06:15 We always loved cinema that tried to break cinema or I tried to open up to a new, not just new forms of storytelling or of visualization, but that questioned the integrity of montage, that questioned the integrity of one task, the kind of assumption of how you get from one image to the other. So in that sense, just as we did with pirate cinema, which was just the screening space, we always said like, no, we're making an invention in cinema, in the field of cinema. Cinema can be the art, or can be the space, or can be the economy, or can be the legal regime.

Pirate cinema was started in 2004, it took place at Bootlab. Both phases of Bootlab between 2000 and 2007 within the same building. We had space for screenings and pirate cinema was kind of a weekly practice. We were involved in any way which included growing an archive of pirated movies. So that turned into a resource.

- JG: 07:20 And at the same time, working on or with Internet-based media as a form of exploring this space of how you can use computer-based databases, websites, but also then moving images in a form that is no longer confined to the idea of this space of cinema. But also then maybe what it was, what is the internet as a medium and to see what happens when these two things kind of come together as a new structure. One thing that we have on Oxdb is this timeline view, which represents the film at a glimpse, get an idea of the structure of the entire film. Instead of watching the film, you can see this timeline and start also to develop a vocabulary and you can start to see things in timelines that allow you to very quickly know how a film is made, or what kind of film it is, or maybe you can recognize the director from the timeline. There are signatures in the temporary composition of a film and the colors they use., But also a full text search and then the ability to extract elements from the film.
- SL: 08:48 It was one of the reasons behind doing Oxdb, because we also didn't want to sit on our own archive. But it seemed very unattractive to just amass a private collection of things. But that we wanted to make them accessible for a large number of people and then in a subsequent step also make our tools for making this and the entire platform accessible. It's very different dealing with 110 films, or hundreds of films, a thousand films or 10,000 and I assume 100,000 is again different. The thing is, these steps are not discreet, you slowly move towards them. One of the things with collecting stuff is that only when you already deep into collecting, you realize that you're collecting. These things emerge over time and they change in shape.
- SL: 09:38 I think probably the moment where you want to start a database, where we thought we have to build something was when [you realize that] you cannot, can no longer watch all of it. I think that's one of the points. Of course you can invent a cinema. Okay, if we can't watch it, maybe everyone else can help us watching things and it forces you yourself. You're not going to just sit down for fun to watch it. You create a certain

rhythm. You make a plan, you make a plan with people, you collectivize it. But I think at some point you realize, okay, some of the watching should be done by machines and a lot of the collecting you do, if you call it collecting, a lot of the archiving you do is something that you do because you know that you first feed it through an algorithmic process. You first feed it through a kind of a pipeline of technical steps, at the end of which are at one point of which, these films end up on Oxdb and that makes this process of collecting a lot more meaningful.

JG: 10:45 The videos are part of the database and we also treat the video material as data. So there is not only metadata about videos, but the videos as it become part of the database and that shows up in certain derivatives of the video. So every frame of video can be addressed. It has a unique URL and there are time-based annotation. So every frame, but also every video clip inside of a film can have the time-based annotation. This could be a subtitle, something that we automatically collect because of they exist already, but it could also be something else. It could be a location that is, then it can also be shown on a map. It could be an event, a historic event that you then can also correlate on a calendar.

JG: 11:40 But it could also be a note or a keyword or a link to another clip or a link to another resource on the Internet. And then you have the ability to search through all of these information that is available. We have multiple timelines by now. So there's the idea of the slit-scan camera, a slit-scan timeline where we have the center pixel of every frame representing with a bit of height information of height, but you only have one frame and there's an timeline that shows the average color of every frame. There's a timeline that has more of a key frame view that is maybe more known from video editing where you see the last key frame and then changes to an the next frame. But you can also just look at the wave form for example, if you're interested in the sound.

SL: 12:35 In some cases they are kind of other representations like posters, and poster frames, and icons for these films that are being generated. Plus all these sub clips in cases where there's time-based annotation so that you'll have these kinds of addressable clips. Metadata will be retrieved from various online sources. We also tried at some point, to add that for each frame. This is more than a billion, I think it's between one and two billion frames in the entire archive. That for each frame we

compute two different hashes, which is kind of the idea of reducing what's in the frame to a 64 bit number. And of a very, very simplistic but interesting approach, or kind of very first step towards image recognition or towards being able to say, okay, show me similar frames, show me things that are visually similar, which is relatively good at finding things like sunsets or horizons or relatively simple shapes.

JG: [13:41](#) So once we had started with pad.ma, which is on another online repository of videos, we realized that we are replicating the same things in different projects. We had ideas for other projects we wanted to build that might use similar things. And so we decided after running Oxdb for three years to rewrite the software that we had in these different projects and create a platform for video archives. And it's a platform is called pan.do/ra. This is an open source software project that we in that allows you to run a film database that has these functionality.

SL: [14:27](#) It's so, it's very paradoxical in the sense that we first did the instance and only later that the framework. But I think it was, it was pretty crucial to do it that way. There was an immediate interest in seeing results in 2007 and to try to get Oxdb done as quickly as possible. And the process of building the foundations for it to making a generalizeable to kind of do it properly, that had to happen later because that was a much more laborious process and required a lot more. We only did it because we knew, okay, we already had more context and more potential projects to work with. That did seem worth spending almost two years on building pan.do/ra

JG: [15:15](#) You can configure various aspects of the platform in that and you can decide what part of the function that you want to use. And this is now also used not only for the sites we run but also for other projects.

SL: [15:30](#) Oxdb is geared at people who use it intensely. it's complex in the sense that if you want to import videos, manipulate video, watch video, annotate video, manage these processes with multiple users. It's not so different from a video editing app in terms of complexity and what it does. And we needed some sort of professional or some sort of consistent working user interface for it, which we couldn't just find elsewhere. So we had to do it ourselves.

- JG: 16:03 When we were starting with these projects online video was largely confined to flash-based video players. This idea of html5 video, that video becomes a part of the open web, that was something that didn't exist, that was just developing. And so we didn't want to build a flash-based, just system, started to use a html5 video. And through that process as we got involved with the people specifying the html specification and have having discussions of what kind of interaction we imagined we wanted to have. We were relatively early working on this online video in that way. And this allowed us to kind of bring in some of our requirements into the specification.
- SL: 16:58 There's a website, it's called pan.do/ra, that's the URL. We're very much into this confusing URL, but we find it kind of quite pretty. So that's where you can get the source code. That's where you can download it. That's where you can also get a pre-installed virtual machine install disc image, where you can just download that and run it without having to install the thing. That's where you find documentation, info, tools around it and various other resources around. So it's a proper, full-fledged open source software project. And it also kind of, if you go to pan.do/ra, you also can see what other people are doing with it.
- SL: 17:48 Oxdb is kind of our own library. There's nothing user contributed directly in it. You cannot upload your own stuff to Oxdb. you can tell us that they think something is missing. But that's all. So it's basically what we find online, what we use what we think is interesting, but there's no curatorial concept. There's a bit of an idea that Oxdb is about. It's about cinema that is relatively hard to find online, that it's not something that's exclusively Hollywood, exclusively western or exclusively contemporary or exclusively well-known stuff, et cetera. But this is relatively hard to say. And this criteria also change from 2000 to 4,000 to 6,000 to 8,000 items. This change, this changes a lot,
- JG: 18:44 But it's also covering the spectrum of moving images. So it's not the only films, but it also contains some video art, some TV series. There are things that are more online games, recordings and things like that. So what is cinema also a question that this kind of database covers.
- SL: 19:19 Obviously as with everything cinema online, the spots that are underrepresented are African cinema, Asian cinema, Southeast Asian cinema, early cinema, off-cinema, a lot of these

contemporary regional cinemas in the south that are kind of defy collection fever because the production is so fast or so ephemeral. So, of course there's are large fields that we don't cover. Part of the ecosystem were prior things we were occupied with textz.com., v2v, other online projects and their communities or context that they provided. Relatively soon, the ecosystem with Oxdb and pan.do/ra grew fastest in India, in Bombay and Bangalore it later grew fastest in Cairo and maybe Beirut and then Turkey.

SL: [20:35](#) So it took an unforeseen trajectory, because these were the places where we found collaborators who are interested in using the software for other purposes than ours. Well, in the case of Pad.ma for example, Shaina Anand and Ashok Sukumaran who run Camp, an artist collective in Bombay, they visited us in Berlin in 2007 and we were actually still working on, on Oxd. When they saw it, they immediately had an idea about it that we hadn't had, which was related to the situation of film making and archiving and internet in India. Concretely, the case that they were working with, or friends with, quite a number of initiatives, individuals or organizations who sat on large amounts of moving images, super interesting materials that were largely inaccessible. That was either not digitized or digitized but not online or kind of abandoned.

SL: [21:43](#) And that they had this idea that not you could not just -- which is the normal argument of archives that you keep something from disappearing and preserve it for the future -- that not only you could do that with a system like Oxdb, that you could actually create an online platform that favored different types of watching than just watching a movie from a to z, pad.ma has a footage archives. So one that would rather make processes of film making, different ways of watching footage, different types of analysis, different types of engaging with materials possible than just this kind of classical cinema, beyond the classical form of cinema, of documentary, of documentation, of political activism, etc. So that was the initial thing that brought us to Bombay.

SL: [22:45](#) And then similarly, funding's probably also part of the ecosystem. So while initially, pad.ma was supported in through a funding structure that was modest, kind of classical Europe-India partnerships, NGO-stuff that usually happens under a banner of development in the broadest sense, and a kind of aid. The over the course, the project has shifted. More recently,

we're more funded through an art kind of context, through an international art funding, which changed the project also because on the one hand it left us with fewer formal bureaucratic requirements, and on the other hand provided us with contacts to a lot of people who were kind of in that same field of critical media practices in the South, in the Arab world, in South Asia, in this whole kind of region between northern Africa and Indonesia.

JG: [23:52](#) There was a collaboration with the Afghan film archive in Kabul where people from India we're going there and found material that they had in the archive but didn't have it digitized. They had a workshop there where they digitized a lot of material that they put on pad.ma and a lot of the filmmakers and cameramen saw the material again after not seeing it for 20 years and then [put] their annotations also on the side. And this entire process then ended up being shown in Documenta in Kassel where you see the pad.ma interface on a screen where we have an automated system that basically uses the website as a form of VJ system where it shows you the interface and then the video plays and annotations are visible, but then it switches to a map for some time and then it goes back to another video, disabled some annotations. So it shows you a text.

SL: [24:55](#) It stands in a tradition of autonomous archives, It's unauthorized. It's self-authorized if you will. it's something that we didn't, we didn't ask for permission. We, there's no official institution behind it. There's no guarantee that we were run it forever, etc etc. But we realized relatively early in our involvement with kind of online archives and online piracy that when we spoke to people who represented official institutions and cinema, for example, official film museums, official institutions, people who had big collections that were totally kind of legitimate in a way, they would tell us usually that, ah, you know, we understand this pirate project of yours very well because that's how we did it, too. All of the film museums, all the museums also, it's not just a matter of film. There are no legitimate museums. There is no legitimate way of starting to collect. It's all stolen stuff.

SL: [25:49](#) You see it with the big museums, European ones, that all have this colonial history. It's very evident they don't legitimately own the things they have. Of course, for Oxdb we had to think about copyright while we were doing it because quite obviously so much is kind of built upon illegal activities and on top of file-

sharing. There are many ways to argue this. I mean we could argue that this is a resource that actually retrieves data from these resources, but it's itself an imaginary archive in the sense that it tries to show, and we always had this idea, hey, maybe it's just fiction. Maybe it's like a movie. Maybe it's something that you can say like, no, this is not an existing database. It's just the kind of it, just imagine this existed and then to be able to do this with an actual website. Imagine what you could do if copyright wouldn't restrict you from taking and recombining and reassembling and representing the history of cinema in this way. But I think where we were quite conscious about what we were not trying to provoke with Oxdb in terms of copyright. And there's several technical aspects by which we tried to keep it in the light gray area and stay able to argue that no, what we're doing is absolutely within the bounds of copyright fair use, research, noncommercial, et cetera, et cetera, et cetera.

- SL: [27:19](#) We always thought we were engaged in the process of creating a new generation of autonomous archives. The ones after the analog to digital transition where it's, on a purely technological basis, much easier to collect. So in a way collecting has been democratized. It's not so much access to materials only, but also access to means of distribution, kind of an understanding of pirate resources as future archives as probably the places that will feed into the future institutions of cinema. We actually were convinced that this kind of combination of computing and networking, concretely bittorrent at the time we started pirate cinema, that this was the most interesting thing in cinema that could have happened. That was much more interesting than anything that happened within cinema. Because it precisely promised to open cinema, not just to new consumers, but that it opened up new ways of producing cinema by organizing distribution differently. Also the spheres of production and consumption would be affected and that it promised, to make possible new types of cinema.
- JG: [28:55](#) Some of these things that we try to do is to bring these things out so that they become maybe not only our own imagination but also something that others can imagine. So I think that is also part of these projects. This is also to shape the discussion and allow for rethinking how technology can be used
- SL: [29:21](#) Our idea from the very beginning was always to be very lightweight. It should cost very little and many of the processes should be automated. So even if we don't have time for it

anymore, it would still run. It could still work this way and if we don't stop watching movies, then it will also grow quite automatically. A lot of these processes just work. The most crucial challenge is not so much does it grow it or how can we maintain it. But how is the software going to look 10 years from now, in an internet that will have changed again.