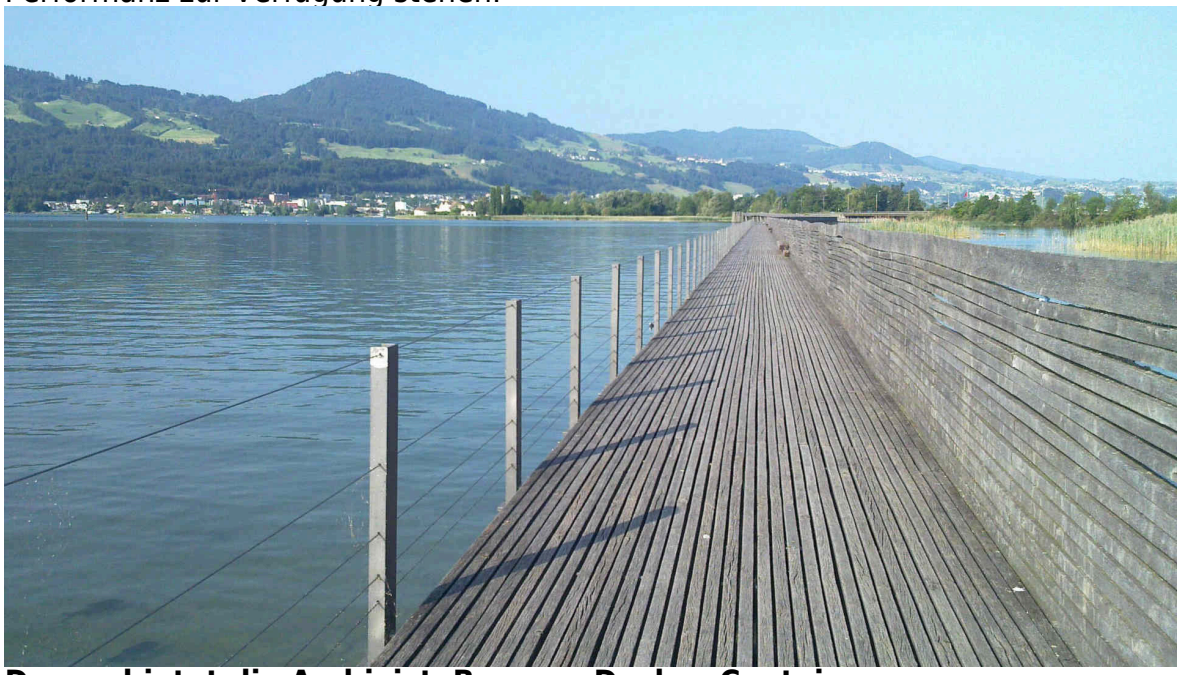
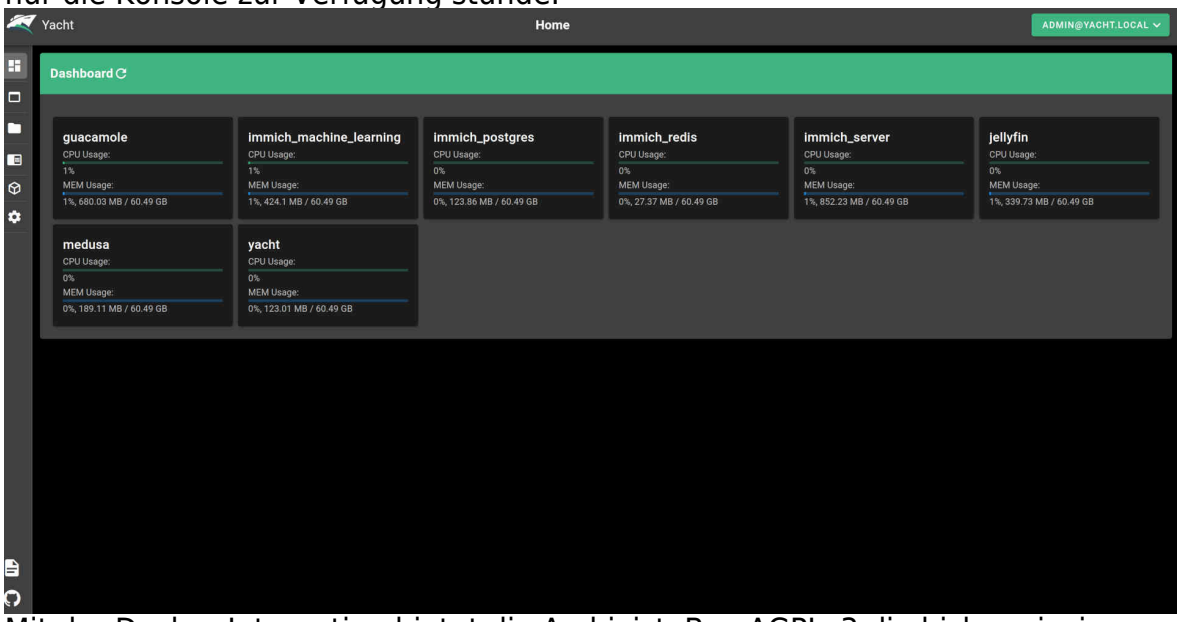
ArchivistaBox 2026/III: Die neue KI-Server-Plattform
Egg, 24. März 2026: Mit dem neuen Release bietet die ArchivistaBox erstmals sämtliche derzeit führenden Server-Technologien im Open Source Umfeld auf einer Plattform an. Neu zu den bisherigen Standbeinen kommen Docker-Container, die lokale KI-Plattform Ollama (lokale künstliche Intelligenz) sowie die derzeit beste Spracherkennung Whisper (inkl. Schweizer Mundart). Möglich macht dies die Integration der ROCm-Technologie, mit der die AMD-Grafikkarten für sämtliche Aufgaben im Bereich der künstlichen Intelligenz in höchster Performanz zur Verfügung stehen.



Darum bietet die ArchivistaBox neu Docker-Container an
Aktuelle Software-Pakete basieren auf vielen Abhängigkeiten. Vereinfacht gesagt kann gesagt werden Software Alpha bedingt Bibliothek Beta, diese aber läuft nicht auf Betriebssystem Gamma, sondern nur auf Delta. Der klassische Ansatz mit der Server-Virtualisierung, der seit mehr als 15 Jahren mit ArchivistaVM (VirtBox) zur Verfügung steht, bietet zwar die Möglichkeit, die Software Alpha mit der Abhängigkeit Beta zu betreiben, indem das System Gamma als virtuelle Instanz aufgesetzt wird. Darauf kann dann Alpha mit Beta installiert werden.
Docker-Container bieten die Möglichkeit, dass Alpha auf Delta läuft, indem Beta und Gamma in einer Art «abgeschlossenem Raum» (Container) zum Laufen erweckt werden. Docker-Container sind für Software-Entwickler in diesem Sinne «lukrativ», weil die eigene Software genau jene Komponenten mitliefern kann, ohne dass ein komplettes Betriebssystem bei allen ausgelieferten Instanzen eingespielt bzw. gepflegt werden muss. Für Anwender/innen ist der Vorteil weniger ersichtlich, denn ohne grafische Oberfläche sind sie oft mit Docker «verloren».
Das Geschäftsmodell von Docker besteht nun darin, eine grafische Oberfläche zur Verfügung zu stellen. Für Private und Firmen unter einem bestimmten Jahresumsatz ist der Einsatz zwar kostenfrei, aber die Lösung ist nicht Open Source. Erstaunlich dabei ist, dass dieser feine Unterschied im Open Source Umfeld kaum Beachtung findet. Und so kam es, wie es kommen musste: Einige nicht ganz nebensächliche Projekte im Open Source Umfeld stehen aktuell nur noch als Docker-Container zur Verfügung.
Natürlich konnte diese Technologie bisher auf jeden ArchivistaBox in einer virtuellen Instanz betrieben werden. Dieser Weg ist aber eben nicht ganz so trivial, zumal dann auch erst Docker auf der Konsole zur Verfügung stünde. Mit

dem Release 2026/III ist Docker daher neu direkt auf der ArchivistaBox GPLv3 vorhanden. Docker wird dabei mit der Web-Applikation Yacht ausgeliefert. Damit gelingt insbesondere «Docker-Neulingen» der Einstieg weit einfacher als wenn nur die Konsole zur Verfügung stünde.



Mit der Docker-Integration bietet die ArchivistaBox AGPLv3 die bisher einzige Open Source Distribution an, welche kostenfrei und ready-to-use sowohl eine webbasierte Verwaltung von KVM/Qemu wie auch nativem Docker an. Dies im Unterschied z.B. zu Proxmox, wo Docker-Container in die LXC-Technologie (Linux-Container) überführt werden (müssen). Der letztere Weg mag für viele Szenarien eine Alternative darstellen, aber letztlich steht dabei Docker nicht nativ zur Verfügung. Dies wiederum führt dazu, dass eben nicht alle Docker-Instanzen laufen.

Native Unterstützung für AMD-Grafikkarten

So leistungsfähig aktuelle Prozessoren auch sind, für gewisse Anwendungen ist eine Grafikkarte (GPU) die bessere Wahl. Dazu ein Beispiel: Bei der Spracherkennung erledigt eine dezidierte Grafikkarte den Job oft um den Faktor 10 oder noch mehr schneller. Nun ist es im Open Source Umfeld so, dass der Marktleader NVidia nur «geschlossene» Treiber ausliefert. AMD stellt hier die bessere Wahl dar, weil die Treiber offengelegt sind (Firmware exklusive). Weiter bieten AMD-Grafikkarten das bessere Preis-/Leistungsverhältnis. Dem gegenüber steht bzw. stand, dass die entsprechenden Softwarepakte (z.B. Blender) oft nur NVidia-Grafikkarten unterstütz(t)en.

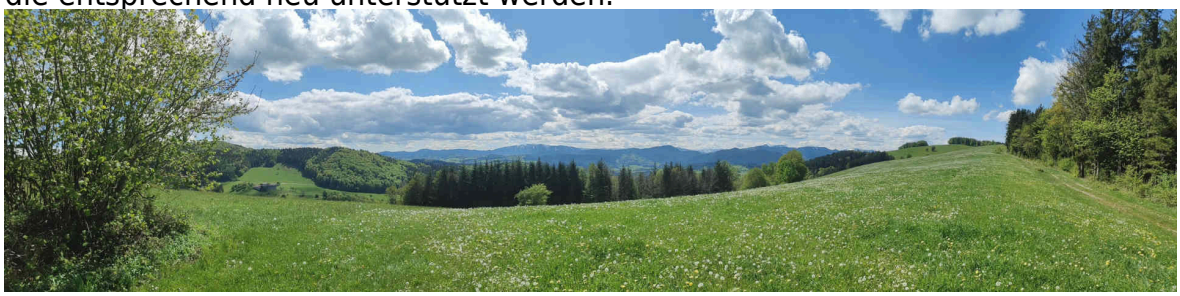
Als 2021 unsere beiden Flaggschiffe ArchivistaBox K1 und Everest auf den Markt erschienen, wurden sie mit AMD Radeon Grafikkarten RX 580 ausgeliefert. Die ROCm-Treiber standen zwar damals zur Verfügung, aber die Software-Integration klappte 2021 oft nicht oder zumindest oft nicht zufriedenstellend. 2026 sieht dies anders aus, ob KI-Modelle, Rendering und/oder Video-Processing, die Treiber-Unterstützung ist aktuell verfügbar. Fairerweise ist hier anzufügen, dass die nahtlose Integration noch immer Probleme darstellen kann. Genau darum gibt es die ArchivistaBox 2026/III, bei der alle Treiber von Haus aus inkludiert sind.

Neue Flaggschiffe K2 und Everest

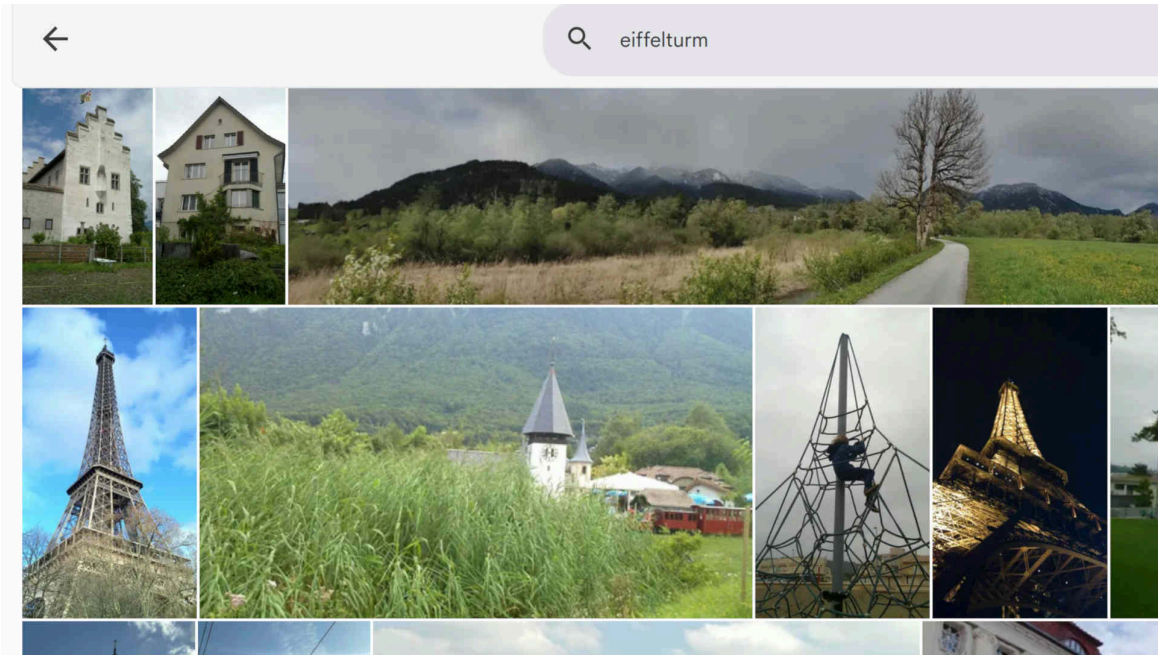
Mit der ArchivistaBox 2026/III bieten wir für die Modelle K1 und Everest ein Upgrade bei der Hardware an. Das Modell K2 bietet von der Leistung her nahezu die Leistung des Modelles Everest aus dem Jahre 2021, ohne dass hier eine eigene Grafikkarte notwendig ist. Beim Modell Everest ist neu die Grafikkarte RX 9070 XT mit 16 GB VRAM mit an Bord.



Damit die RX 9070 XT (und andere AMD-Grafikkarten) angesprochen werden können, sind aktuelle Firmware-Treiber bzw. Linux-Kernel notwendig. Daher steht neu bei der ArchivistaBox 2026/III der Linux-Kernel 6.18.18 zur Verfügung. Die Linux-Firmware stammt von Mitte März 2026. Auch diese musste integriert werden. Als kleine «Nebenbemerkung» sei angefügt, dass die aktuelle Firmware für alle unter Linux verfügbaren Geräte 1.8 GByte umfassen. Vor etwa drei Jahren waren es noch ca. 0.9 GByte bzw. die Hälfte.
Dies führte dazu, dass bei der ArchivistaBox auch der Boot-Mechanismus angepasst werden musste. Bisher wurde die gesamte ISO-Datei in die Boot-Partition (5.6 GByte) abgelegt. Da bei einem Update die Daten kurzfristig zweifach vorliegen, hätte das Release 2023/III die Boot-Partition «gesprengt». Neu besteht ein kleineres essentielles Image (ca. 1.7 GByte), womit die bisherige Grösse der Boot-Partition für die nächsten Jahre «gerettet» ist.
Nachdem all diese Hürden gemeistert waren, konnte die ArchivistaBox endlich mit der Grafikkarte RX 9070 XT gestartet werden. Selbstverständlich profitieren von diesen Anpassungen auch viele andere aktuelle Hardware-Komponenten, die entsprechend neu unterstützt werden.

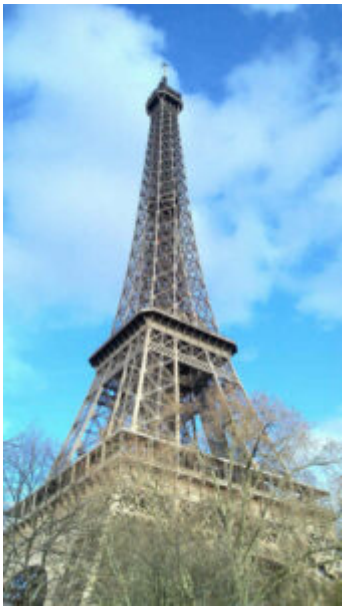


22 GByte: Die ROCm-«Treiber» von AMD
Mit dem Upgrade des Linux-Kernels auf die Version 6.18.18 war das Fundament gelegt. Damit moderne Grafikkarten für Render- und KI-Aufgaben gerüstet sind, werden zusätzliche Treiber benötigt. Wobei Treiber jetzt nicht primär der richtige Begriff wäre. Im Falle von AMD ist es das ROCm-Paket. Es umfasst aktuell ca. 22 GByte. Richtig gelesen, 22 GByte auf der Festplatte sind notwendig, damit die Grafikkarte für Aufgaben im Bereich der Grafik bzw. der künstlichen Intelligenz überhaupt genutzt werden kann.
Zum Vergleich sei gesagt, die aktuelle ISO-Datei der ArchivistaBox benötigt ca. 4.1 GByte. Somit fallen für die Grafikkarte um den Faktor 5.x mehr Daten an. Natürlich darf die Frage gestellt werden, ob weniger nicht mehr wäre? Sicher wäre weniger mehr, aber letztlich wäre hier der Anbieter AMD gefordert, weniger auszuliefern. Oder Mitbewerber/innen, welche entsprechende Produkte auf den Markt bringen. Oder Software, die mit weniger Leistung klarkommen würde.
Lokale KI mit Ollama
Nachdem die ROCm-Treiber von AMD korrekt integriert waren, konnte die erste lokale KI gestartet werden. Oder besser gesagt, die erste KI wurde natürlich längst vor der ROCm-Integration aufgesetzt, um danach festzustellen, wie «grottenlangsam» die CPU unter Höchstlast überhaupt irgendwann Ergebnisse liefert. Als Beispiel sei die Bildkategorisierung mit dem neu in der ArchivistaBox verwendeten Modell angeführt. Pro Bild werden hier mit dem Prozessor Ryzen 9590 3XD unter Verwendung aller 32 Threads ca. 2 Minuten benötigt. Über die Grafikkarte mit ROCm und der RX 9070 XT sind es ca 1.1 Sekunden.



Beispiel einer mangelhaften Bilderkennung (Immich) mit dem Suchbegriff ‚Eiffelturm'

Natürlich kann auch eine Software (wie z.B. bei Immich) verwendet werden, die deutlich schneller bzw. ohne Grafikkarte arbeitet, um Bilder zu erkennen. Allerdings entsprechen die Ergebnisse nicht dem, was State-of-the-Art ist. Dazu sei das Suchbeispiel «Eiffelturm» angeführt. Obenstehend findet sich eine Trefferansicht mit Immich. Zwar ist es nicht so, dass der Eiffelturm gar nicht gefunden wird, allerdings laufen die korrekten Treffer unter ferner liefen.

Zum Vergleich dazu, die neu vorhandene Engine der ArchivistaBox erfasst das Bild links mit: **«Spielplatz; Klettergerüst; Kind; Klettern; Drahtnetz».** Das Bild rechts, der «richtige» Eiffelturm wird mit **«Eiffelturm; Paris; Frankreich; Stahlkonstruktion; Himmel»** beschlagwortet. Das sind derart gewaltige Unterschiede, dass die ArchivistaBox hier doch lieber eine gute Engine «einbaut», auch wenn dazu eine native Grafikkarte notwendig ist.
Die Bilderkennung bildet nur eine spezifische Anwendung von Ollama. Aktuell stehen einige hundert Modelle zur Verfügung, die bei der ArchivistaBox 2026/III alle ausschliesslich lokal laufen (bei Ollama muss der Netzzugriff deaktiviert werden). Dabei gibt es die allseits bekannten Chats, aber auch Übersetzungen sowie Modelle für die Inhaltszusammenfassung. Die Ergebnisse sind dabei schon «verdammt» gut, um es etwas salopp zu sagen.

Spracherkennung mit Whisper

Bereits seit 2021 verfügt die ArchivistaBox über eine Spracherkennung. Die damaligen Tests mit Schweizerdeutsch erzeugten nicht wirklich brauchbare Ergebnisse. Als Beispiel sei hier eine kurze Sequenz aus dem Film «Die Schweizermacher» (Rolf Lyssy, 1978, ab ca. 50 Sekunden bis 1 Minute 20) angeführt, zunächst das (bisherige) Ergebnis mit Vosk:

59.79: lange sie mir vorbeigeschaut habt
62.94: wo mir wie man der ausländer berufsbild sanieren
66.57: damit der bürger äußerem land mehr erfahren
71.07: wir mussten sie neutral riesig für soli solid zum wieviel leide rock gebilligt wird

Nun die aktuelle Ausgabe mit Whisper:

00:00:57: Nennen Sie mir ein paar Eigenschaften,
00:01:02: die wir bei einem Ausländer voraussetzen müssen,
00:01:06: damit er Bürger von unserem Land werden kann.
00:01:10: Wie muss er sein?
00:01:11: Neutral.
00:01:14: Fleissig.

Unschwer lässt sich erkennen, dass die «moderne» Engine selbst bei Schweizerdeutsch nicht nur fast 100% eine korrekte Übersetzung ins Hochdeutsche hinkriegt, sondern auch den Satz so wiedergibt, wie er gesprochen wurde. Eine solche Leistung mit Open Source zu realisieren, doch dies ergibt (nach viel viel Arbeit und Ungewissheit) doch auch ein gutes Stück Stolz. In diesem Sinne kann angeführt werden: Freude herrscht. Für unsere bisherigen, für neue Kunden und die Open Source Gemeinschaft.
Zu Whisper noch zwei Anmerkungen: Wo viel Licht ist, gibt es auch etwas

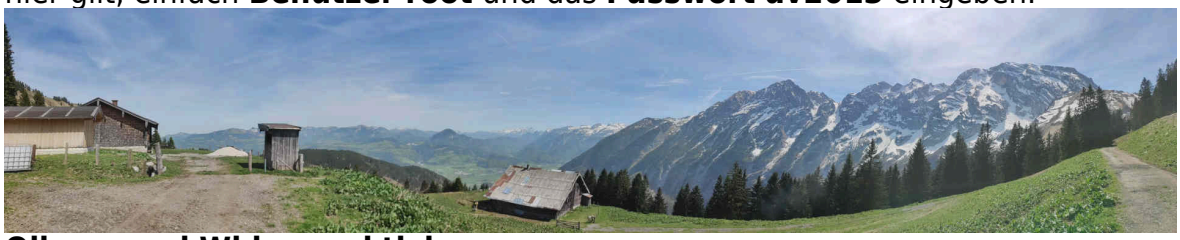
Schatten. Whisper (wohl passend zum Namen) hat zuweilen die Angewohnheit, wo nichts ist, einen Text zu «halluzinieren», d.h. manchmal, wenn der Ton gegen Null tendiert, (er)findet die Engine Text, wo keiner ist. Ferner sei angefügt, dass spezifisch für die Schweizer Mundart trainierte Modelle eine schlechtere Qualität lieferten als dies bei der neusten Version von Whisper der Fall ist.

Digitale Souveränität mit der ArchivistaBox

Die ArchivistaBox steht seit ca. 20 Jahren dafür da, dass unsere Kunden ihre Daten lokal halten. Die ArchivistaBox bietet dazu fixfertige Lösungen an, die mit der Version 2026/III neu auch für die Bereiche Container-Management (Docker), lokale KI-Agenten und Bild- wie Spracherkennung (Whisper) als 100 Prozent Open Source Lösung zur Verfügung stellen. Wer aktuell im Business-Umfeld seine Lösungen in die Cloud stellt, verpasst für längere Zeit und nachhaltig die Chance, digital souverän agieren zu können. Im aktuellen politischen Umfeld erscheint dies eine gewagte Wette.

Dabei wird oft argumentiert, es stehe keine lokale Alternative zur Verfügung. Die ArchivistaBox 2026/III zeigt eindrücklich, dass dem nicht so ist. Die ArchivistaBox kann ready-to-use mit Wartung und Support bestellt werden. Wer dies lieber selber machen möchte, kann die **ISO-Datei (2026-er Release) beziehen und selber installieren.** Die ArchivistaBox bietet dabei eine radikal einfache Lösung für das Self-Hosting an, die auch Open Source ist.

Docker z.B. kann nach dem Hochfahren der ArchivistaBox (ca. 10-20 Sekunden) direkt über den gleichnamigen Link gestartet werden. Einfach bei **Benutzer admin@yacht.lokal** und **Passwort av2013** eingeben, und los geht es. Das gleiche gilt natürlich auch für die bisherige Server-Virtualisierung (VirtBox). Auch hier gilt, einfach **Benutzer root** und das **Passwort av2013** eingeben.



Ollama und Whisper aktivieren

Für die Komponenten Ollama wie auch Whisper wird eine moderne AMD-Grafikkarte empfohlen. Weiter ist die Datei ai.os zu beziehen. Sie umfasst ca. 30 GByte, beinhaltet dafür sämtliche Komponenten, wie sie eben benötigt werden. **Die Datei ‚ai.os' (30 GByte!) muss auf der ArchivistaBox im Verzeichnis /home/archivista/data vorliegen.** Nun über das Terminal mit 'su' zum Benutzer root (Passwort av2013) wechseln und die Datei entpacken: ‚unsquashfs -d ./ai ai.os'. Danach die Engine mit ‚cp ./ai/desktop.sh /home/data/archivista/cust/desktop' aktivieren und einen Neustart durchführen. Nach dem Neustart kann im Terminal ‚ollama list' aufgerufen werden.

```
NAME ID SIZE MODIFIED
xyz44:latest b3cacc3eb7f 3.3 GB 18 hours ago
```

Selbstverständlich können danach weitere Modelle hinzugefügt werden. Die Whisper-Engine kann über das Skript ‚mp2txt.sh' im Verzeichnis ‚/home/archivista/data/ai' aufgerufen werden. Aktuell gibt es weder für Whisper noch für Ollama eine grafische Oberfläche. Eine entsprechende Oberfläche darf in einer der nächsten Versionen der ArchivistaBox erwartet werden.

Sowohl Whisper wie Ollama laufen auch ohne eine eigene AMD-Grafikkarte. Allerdings ist die Geschwindigkeit selbst bei einer schnellen CPU um den Faktor 1:10 bis 1:50 langsamer. Kurz, es lohnt sich vielleicht für einen Test, nicht aber für die Produktion. NVidia-Grafikkarten werden aktuell nicht unterstützt. Die Community ist herzlich eingeladen, diesen Part zu «vollbringen».

Preisbeispiele bei ArchivistaBox K2 und Everest

Für Business-Kunden stehen die Modelle K2 ab 2990.– bzw. Everest ab 4990.–

zur Verfügung. Diese Preisbeispiele entsprechen den aktuellen Marktgefüge bei den SSD-Platten bzw. Speicherbausteinen (RAM). In diesem Sinne können die entsprechenden ArchivistaBoxen nur auf sehr kurzfristige Zusicherung bestellt bzw. ausgeliefert werden. Die Hardware ist gemessen an der Leistung sehr kompakt (30x33x20cm) und passt daher nicht nur in Server-Umgebungen, sondern auch unter jeden Schreibtisch.
Im durchschnittlichen Büro-Alltag arbeiten sowohl K2 als auch Everest flüsterleise. Die neuen Flagschiffe können in Egg jederzeit an einem Open Friday «beschnuppert» werden. Die gesamte Software der ArchivistaBox (inkl. KI-Module) können frei bezogen und betrieben werden. Support und Wartungsverträge sind sowohl für die Hardware als auch für die reine Software der ArchivistaBox verfügbar.
P.S: Falls hier jemand Informationen zu unserem Flagschiff Dokumenten Management System bzw. ArchivistaDMS vermisst, dem sei gesagt, dass die hier eingeführten Basis-Technologien in Zukunft nahtlos in ArchivistaDMS einfliessen werden. Als Beispiel seien die obigen Beispiele Bild-Klassifikation und Spracherkennung angeführt.