***3.2.1 What are the shortcomings of our very simple tagging approach?***

Unser Ansatz geht nur auf das Containerformat ein. Informationen können aus der Quelldatei immer hergestellt werden. Erst eigene Tags verknüpfen die Dateien mit neuen Informationen.

Die Qualität des Systems steht und fällt natürlich mit der Anzahl der gesetzten Schlagworte bzw. Tags. Nachdem der hier verwendete Ansatz relativ simpel gehalten ist, sind die Möglichkeiten vergleichsweise begrenzt. Eine einfache Verbesserung würde es etwa schon sein, wenn nach mehreren Tags gleichzeitig gesucht werden könnte.

***3.2.2 What can be said about relationships between tags?***

Die Tags sind nicht miteinander verknüpft. Jedoch kann eine Datei mehreren Tags zugeordnet sein und Tags mehreren Dateien. Sie gruppieren die Dateien untereinander, wenn diese die gleichen Tags enthalten. Der Algorithmus aus 3.3 vergibt Tags nach Farb-Schwellenwerten und sorgt so für eine zusätzliche Gruppierung anhand der Bildinformationen.

***3.2.3 We used tags for simple "content typing". How can content typing in the sense we did it improve navigation and search in an information system?***

Es kann gezielt nach Inhalten gesucht werden, in dem die entsprechenden Tags wählt. Ungewollte Inhalte können ausgeschlossen werden und man gelangt schrittweise zu einer Auswahl. Die Methode funktioniert gut für eine schnelle Vorauswahl. Etwa, wenn man nur Bilddateien betrachten möchte und durch einen einzigen Klick bereits alle andersartigen Dateien entfernt hat. Das ist für die Benzuterfreundlichkeit eines solchen *information systems* natürlich nicht unerheblich.

***3.3.1 discuss the advantages and limitations of your heuristics***

Jeder Pixel wird anhand der jeweiligen RGB-Werte betrachtet. Zunächst wird festgestellt, wie viele schwarze (alle drei RGB-Werte sehr niedrig) und weiße (alle drei RGB-Werte sehr hoch) Pixel im Bild vorhanden sind. Diese werden dann von der Gesamtanzahl der Pixel abgezogen. Somit können Bilder mit großen Flächen an Weiß bzw. Schwarz (Hintergrund) betrachtet werden. Sofern alle verbliebenen Pixel in einem der drei Farbbereiche eine signifikante Häufigkeit in Bezug auf die Gesamtpixelanzahl aufweisen (hier wurde 20% gewählt), wird der entspechende Tag hinzugefügt. Ob eine Farbe durch einen Pixel repräsentiert wird, wird dadurch festgestellt, ob der entsprechende Farbwert größer ist, als von beiden anderen Farbwerten zusammen.

Eine Limitierung ist bzgl. der Farbintensität erkennbar: So werden Bilder, die offensichtlich einer Farbe entsprechen, nicht dahingehend getaggt, wenn die Farbintensität eher schwach ausgeprägt ist. Auch Bilder mit sehr dunklen Farben werden hier benachteiligt.

***3.3.2 discuss the advantages and limitations of our simple dominant color model***

Der Algorithmus ist – wie gefordert – simpel gehalten, was den Vorteil ergibt, dass er schnell ist und vergleichsweise verständlich bleibt.

Es sind allerdings nur Kategorisierungen anhand der Farben Rot, Blau und Grün möglich. Als weiteren Nachteil könnte man ansehen, dass die Schwellenwerte, ab derer eine dominante Farbe festgestellt wird (hier ab 20%), nicht variabel sind.

***3.4.1 What (type of) layout algorithm do we use for the information visualization in this assignment?***

Der in GraphML verwendete Graph ist ein Baum. Das Tag Element ist die Wurzel und die Dateien die Blätter. Zwischen den Knoten gibt es Kanten zu weiteren Tags.

***3.4.2 What are advantages/disadvantages of the used infovis?***

Infovis bietet einen direkten Zugriff auf die Dateien und Tags. Es Visualisiert die Zusammenhänge und den Dateibestand im System und erweitert sich bei entsprechender Implementierung selbstständig. Bei vielen Dateien wird es jedoch leicht unübersichtlich und die vielen Elemente überlagern sich teilweise. Die Steuerung ist ebenfalls verbesserungswürdig. Die Tatsache, dass es ein Java Applet ist, das von vielen Browsern geblockt wird, ist ein zusätzlicher Nachteil.