

# Wissensentwicklung mit „Crowd OSInfo“

Eine Innovation des Cyber Documentation &  
Research Center (CDRC) der Zentraldokumentation  
(ZentDok), Landesverteidigungsakademie (LVAk)

Wissensentwicklung mit „Crowd OSInfo“

ISBN: 978-3-902944-80-1



25/15

Klaus Mak, Joachim Klerx, Hans Christian Pilles, Johannes Göllner

Mak, Klerx, Pilles, Göllner

Schriftenreihe der  
Landesverteidigungsakademie



Schriftenreihe der  
Landesverteidigungsakademie

Klaus Mak, Joachim Klerx, Hans Christian Pilles, Johannes Göllner

# **Wissensentwicklung mit „Crowd OSInfo“**

**Eine Innovation des Cyber Documentation & Research Center  
(CDRC) der Zentraldokumentation (ZentDok),  
Landesverteidigungsakademie (LVAk)**

**25/2015**

Wien, September 2015

**Impressum:**

Medieninhaber, Herausgeber, Hersteller:

Republik Österreich / Bundesministerium für Landesverteidigung und Sport  
Rossauer Lände 1  
1090 Wien

Redaktion:

Landesverteidigungsakademie  
Zentraldokumentation und Information  
Stiftgasse 2a  
1070 Wien

Schriftenreihe der Landesverteidigungsakademie

Copyright:

© Republik Österreich / Bundesministerium für Landesverteidigung und Sport  
Alle Rechte vorbehalten

September 2015  
ISBN 978-3-902944-80-1

Druck:

HDruckZ-ASt Stift xxxx/15  
Stiftgasse 2a  
1070 Wien

## Inhaltsverzeichnis

1	Einleitung.....	5
2	Einbettung des CDRC in die Organisation der Zentraldokumentation.....	7
2.1	Zielsetzungen CDRC/LVAk.....	9
2.2	Prozess-Modell CDRC/LVAk .....	10
3	CDRC/ZentDok – Hauptaufgaben und korrespondierende Produkte .....	14
3.1	Identifizieren von OSInfo mit „Relevanz Monitoring“ .....	16
3.2	Dokumentation und Translation von mehrsprachigem Content.....	17
3.3	Fachinformation .....	19
3.4	Forschung und Entwicklung.....	21
4	CDRC Innovation: Crowd OSInfo mit CROWD.....	23
4.1	CROWD – Assoziationsheuristik .....	26
4.2	ZentDok – Fragenheuristik.....	28
4.3	CROWD – Recherche .....	32
4.3.1	CROWD Recherche – Musterprozess .....	33
4.3.2	High Profile Sites (HPS).....	39
4.4	Cyber Berichte des CDRC (6W).....	41
4.5	CROWD – Assoziations-Matrix-Prozess („Chance – Risiko“) ...	42
4.6	Visualisierung der CROWD – Assoziations-Matrix.....	44
5	Zusammenführung von ZentDok-Experten- & CROWD- Rechercheprozess.....	46
5.1	CDRC Innovation: Modell der zeitlichen Quelleninterdependenz zur Informationsverdichtung.....	50
5.2	CYBER Informationsplattformen .....	56

5.2.1	CYBER – Kategoriensystem .....	58
5.2.2	CYBER - DocType, Site Info, Meta Data, Expert Tagging.....	60
5.2.3	News Erfassungssystem LAMARR der CDRC Cyber Informationsplattform .....	62
5.2.4	Recherche Datenbank.....	64
5.2.5	Multi Media Demonstrator .....	65
5.2.6	Terminologiesystem – ProTerm.....	66
6	Horizon Scanning Center „CYBER“ .....	69
7	Ausblick .....	72
8	Literatur.....	75
9	Abbildungsverzeichnis.....	77
10	Autoren .....	79

# 1 Einleitung

Durch die Digitalisierung ist die Dokumentation dem wohl größten Strukturwandel seit der Erfindung des Buchdruckes unterworfen. Häufig wurde dies in letzter Zeit als selbstverständlich und bereits umgesetzt abgetan. Diese Publikation folgt einer anderen Argumentationslinie. Die jetzigen Ansätze zur Digitalisierung sind vermutlich die ersten Schritte zu einer nachhaltigen Transformation, welche alle wissensintensiven Domains betrifft. Sowohl Wissenschaft, als auch Technologieentwicklung und staatliche Administration sind in ihrem Kern abhängig von einem effizienten Umgang mit Wissen. Neue militärische Domains, wie z.B. Cyber, aber auch neue militärisch relevante Technologiefelder, wie z.B. Robotik, künstliche Intelligenz und autonome Systeme, sind ganz entscheidend von Wissen geprägt. Es ist nicht unwahrscheinlich, dass Wissen in zukünftigen Konflikten die entscheidende strategische Ressource wird. In der Wissenschaft und in der innovativen Technologieentwicklung war Wissen seit jeher die strategische Ressource. Daraus ergibt sich ein Kooperationspotential bei den Methoden der Wissensbeschaffung, der Recherche und der Dokumentation, welches im Rahmen des „Cyber Documentation & Research Centers“ (kurz: CDRC) zu ersten Innovationen geführt hat.

Sowohl in der Wissenschaft, der Technologieentwicklung, als auch bei der Informationsbeschaffung für staatliche Organisationen hat sich eine anlassbezogene Recherche durchgesetzt, die zwar schnell zu neuen Informationen führt, aber immer Informationen aus der Perspektive des jeweiligen Analytikers oder Wissenschaftlers liefert. Es drängt sich daher die Frage auf, wie lässt sich die Recherche objektivieren, vor dem Hintergrund, dass Wissen so gut wie immer in unterschiedlichen Sprachen, mit unterschiedlichem kulturellen Hintergrund und mit den digitalen Medien zunehmend auch in den unterschiedlichsten Formaten (Text, Sound, Video) und auf den unterschiedlichsten Plattformen (Pressedatenbanken, wissenschaftlichen Datenbanken, aber auch YouTube, Blogs, Diskussionsforen, usw.) veröffentlicht wird. Der Einsatz einer größeren Menge an Personen („Crowd“) für die Recherche liegt damit auf der Hand. Daraus entstehen jedoch Fragen, wie: Wie viele Personen liefern ein objektives Ergebnis? Wie gut müssen diese Personen ausgebildet sein? Wie lassen sich diese Ergebnisse zusammen führen? Sind für

wissenschaftlich relevante Recherchen Wissenschaftler notwendig oder ersetzt die Vielfalt der unterschiedlichen Perspektiven das Fachwissen eines Wissenschaftlers oder Analytikers? Aus diesen Fragestellungen heraus ist das Konzept der „Crowd Open Source Information“ (Crowd OSInfo) entstanden und im Rahmen des CDRC ausgetestet worden.

Aufbau und Arbeitsweise sowie konkrete Auswirkungen des Pilotprojektes CDRC sind Inhalt dieser Publikation. Sowohl das CDRC als auch seine Einbettung in die ZentDok werden in dieser Publikation vorgestellt und zentrale Innovationen des CDRC werden präsentiert.

Mit Einnahme des Organisationsplanes „OrgPlan 2010“ wurde neben dem Hauptreferat „Dokumentation“ und „Fachinformation“ das Hauptreferat „Wissensmanagement“ erfolgreich in die Zentraldokumentation integriert. Die Hauptaufgaben der Dokumentation und Fachinformation befinden sich im Strukturwandel und wurden und werden in rasantem Tempo in jeder Hinsicht (sowohl qualitativ als auch quantitativ) an die neuen Herausforderungen angepasst. So sind neben dem Mengenproblem an offenen Fachinformationen („Open Source Info“) im Internet vor allem auch Sprachprobleme und Qualitätsprobleme sowie Organisationsprobleme neu zu bewerten und Maßnahmen für eine bessere Verarbeitung und zu Verfügungsstellung für die Nutzer der ZentDok zu entwickeln.

Durch die Möglichkeit Rekruten mit technischen Fachkenntnissen und Sprachkenntnissen im Rahmen eines Pilotprojektes in der wichtigen „Cyber“ Domain für „Open Source Informationsarbeit“ einzusetzen, konnten die oben angesprochenen Fragestellungen adressiert werden. Aufbauend auf den Forschungserfahrungen der ZentDok, wurde ein entsprechendes Konzept entwickelt, von den Fachabteilungen geprüft, vom BMLVS genehmigt und von der ZentDok in den letzten 12 Monaten in zwei Phasen erfolgreich umgesetzt. Einige der Ergebnisse werden in dieser Publikation präsentiert.

## 2 Einbettung des CDRC in die Organisation der Zentraldokumentation

Die ZentDok ist seit 1968 für die Auswertung, Dokumentation und Bereitstellung von offenen Fachinformationen für alle Dienststellen des ÖBH verantwortlich. Dies umfasst neben dem Aufbau und der Pflege von internen Datenbanken auch die Generierung eines Fachthesaurus sowie die Mitentwicklung an allen notwendigen Werkzeugen und Methoden. Besonders die Integration von modernen Suchtechnologien und eines des modernsten Wissensmanagementwerkzeuges seien hier erwähnt. Voraussetzung dafür war die Integration von Forschungsaufgaben in die Organisation, die ein ständiges Innovationsmanagement bis zur realen Umsetzung von neuen Ideen in das Tagesgeschäft ermöglicht<sup>1</sup>.

Im nachfolgenden Diagramm ist die derzeitige Organisation dargestellt und zukünftige Neuerungen grün gekennzeichnet. Die Umsetzungsmaßnahmen, die im Projekt getestet wurden, sind bereits eingeleitet.

---

<sup>1</sup> Vgl. Mak, K., Woitsch, R., „Der Einsatz des prozessorientierten Wissensmanagementwerkzeuges PROMOTE® in der Zentraldokumentation der Landesverteidigungsakademie“, Schriftenreihe der Landesverteidigungsakademie, 19/2005, S.5ff



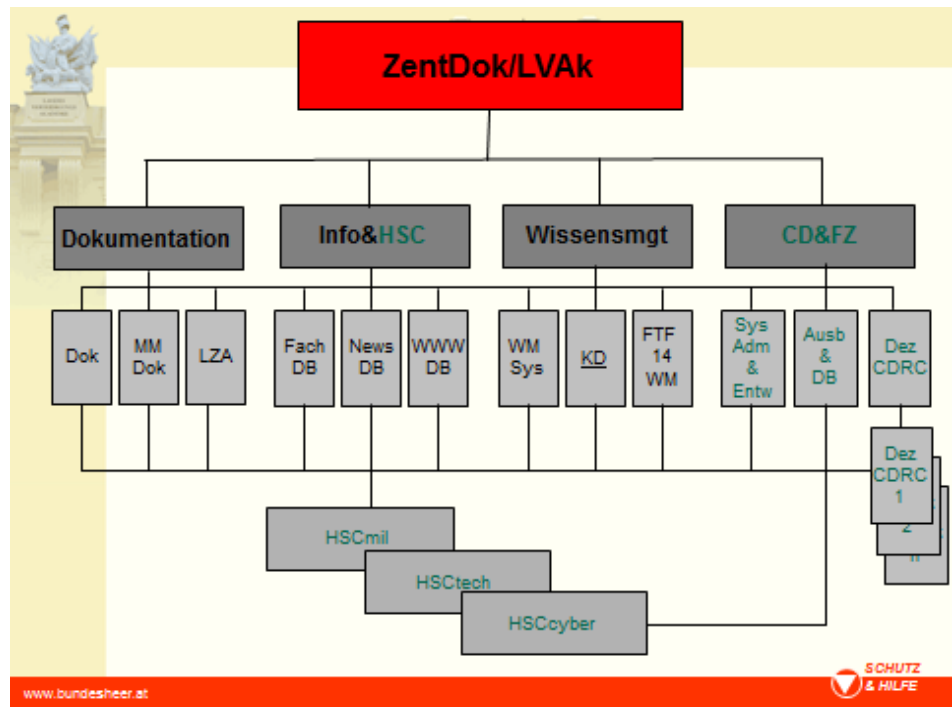


Abbildung 1: Positionierung des CDRC im Organigramm der ZentDok/LVAk<sup>2</sup>

Das vierte Hauptreferat „Cyber Dokumentations- und Forschungszentrum“ („CD&FZ“) bildet mit dem Kaderpersonal und 14 „Cyber Rekruten“ je Einrückungstermin den Ausgangspunkt für die Bewältigung neuer „Open Source Dokumentations-, Informations- und Forschungsarbeit“ bis hin zur Ausgestaltung von sog. „Horizon Scanning Centern“ („HSC“) zu speziellen militärisch relevanten Themenbereichen. Dabei werden in Zukunft auch neue dezentrale Organisationselemente („dezCDRC“) und alle bestehenden Elemente der ZentDok mit ihrer Expertise in die aktuelle Neuausrichtung der Produkte einbezogen.

<sup>2</sup> Quelle: eigene Darstellung

Die Bezeichnung „Cyber Documentation & Research Center“ („CDRC“) - der Projekttitel - wird in dieser Publikation bis zur tatsächlichen Umsetzung der organisatorischen Maßnahmen verwendet.

## **2.1 Zielsetzungen CDRC/LVAk**

Unter dem Motto „Cyber ist nicht nur Risiko, sondern auch Chance!“ sollen in Zusammenarbeit mit Rekruten ihre Kompetenzen erkannt werden und ihre Fertigkeiten, Fähigkeiten, Wissen und Denkweise genutzt werden.

Hauptziele sind „Open Source Informationen“ („OSInfo“) der Domain „Cyber“ unter Mithilfe von „Cyber-Rekruten“ für das ÖBH, staatliche Organisationen und eventuell Partnerorganisationen nutzbar zu machen sowie Forschungs- und Entwicklungsaufgaben für die notwendigen Verarbeitungsmechanismen von OSInfo zu bewältigen.

Ebenso sollte bei geforderten Kompetenzen die nicht vorhanden sind, den Rekruten eine Ausbildung zu teil werden. Durch alle Maßnahmen sollte nicht nur während des Grundwehrdienstes eine Nutzung für das ÖBH möglich werden, sondern gleichzeitig ein Assessment für weitere Aufgaben im Cyber Bereich Österreichs erreicht werden. Neben speziellen Aufgaben im ÖBH, wären dies auch Aufgaben die das SKKM betreffen, aber auch für F&E-Aufgaben oder Bereiche wirtschaftlicher Unternehmen, vor allem kritische Infrastrukturen betreffend, die von der Expertise der Cyber-Rekruten profitieren könnten.

Der Aufbau eines „Expertennetzwerkes Cyber“ wie auch einer „Cyber Miliz“, ist mittelfristig gesamtstaatlich dadurch umsetzbar.

Die Erweiterung der Sprachkompetenzen sowie der Technischen Kompetenzen für das ÖBH wurde während des Pilotprojektes klar ersichtlich. Aber auch im Erkennen gesellschaftlicher, sicherheitspolitischer und militärischer Dimensionen, wurden aus dem „Integrationsprojekt“ CDRC neue Erkenntnisse gezogen. Mit der Erstellung von Heuristiken werden diese Erkenntnisse formalisiert und für das System nutzbar dokumentiert.

Durch die Nutzung von Kooperationspartnern des ÖBH, insbesondere durch den Einsatz eines Gastforschers des AIT, konnten während des

Projektes auch der Nachweis der wissenschaftlichen Qualität der Arbeit erbracht werden und F&E-Aufgaben der Rekruten gemeinsam entwickelt und zu Produktreife gebracht werden. Kooperationspartner aus KIRAS-Projekten, aus Industrie und Universitäten unterstützen durch theoretische und praktische State of the Art Komponenten den Versuch, die Dynamik der Domain Cyber zu veranschaulichen. Auch die Möglichkeit von Testungen und Weiterentwicklungen von Systemen, bis zu vollständig neuen wissenschaftlichen Erkenntnissen und Entwicklungen wurden im Projekt herausgearbeitet und teilweise auch in den Tagesablauf integriert.

Letztendlich gelang es die Grundlage für eine Wissenslogistik in der Domain Cyber für OSInfo zu modellieren und umzusetzen, die in weiterer Folge bereits ein wesentliches Element für ein „HSC Cyber“ darstellt und auch für weitere Domains in Zukunft anwendbar wird.

## **2.2 Prozess-Modell CDRC/LVAk**

Im nachfolgenden Modell wird veranschaulicht, wie das CDRC in das System ZentDok integriert wurde. Alle Aufträge für „Fachinformationsbeschaffung“ oder „Forschungs- und Entwicklungsaufgaben“ von berechtigten Nutzern, werden in einem sogenannten Leitung & Management Board (Ltg & Mgmt Board) koordiniert und den zuständigen Bearbeitungsteams zugewiesen. Mittels Informationssystemen und den jeweils notwendigen Berechtigungssystemen werden alle Produkte innerhalb und außerhalb des ÖBH den Bedarfsträgern zur Verfügung gestellt.

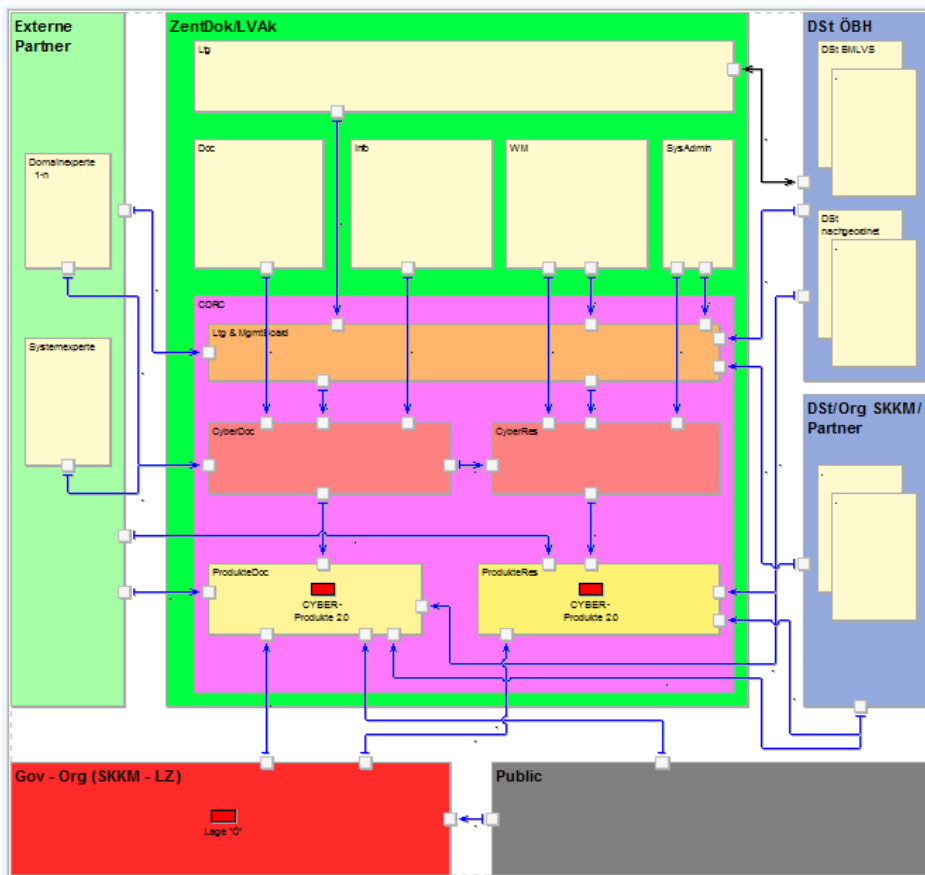


Abbildung 2: Integration des CDRC in die Prozessstruktur der ZentDok<sup>3</sup>

Der zentrale Prozess der Wissensentwicklung im CDRC folgt dabei den Schritten des entwickelten Z-Modells (Knowledge Development Mechanismus für Horizon Scanning Center), zur Erstellung von multidimensionalen Lagebildern, als Grundlage für das Generieren von Strategischen Handlungsoptionen. Wie aus der folgenden Abbildung ersichtlich, beginnt jede Erstellung eines Lagebildes mit der Informationsbeschaffung, für die, vor allem bei wiederholter Anwendung, eine

<sup>3</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok

Informationslogistik aufgebaut werden muss, die zur Generierung eines Informationslagebildes führt<sup>4</sup>.

Der Wissensentwicklungsmechanismus des CDRC gliedert sich dementsprechend in den Wissenslogistik-Prozesscluster 1: Informationslogistik, den Wissenslogistik-Prozesscluster 2: Szenario Planung und -management, den Wissenslogistik-Prozesscluster 3: Risikobewertung und -mapping sowie den Wissenslogistik-Prozesscluster 4: Multidimensionales Lagebild und Decision Making; wobei im CDRC zunächst einmal nur die Informationslogistik im Detail ausdifferenziert wurde. Wie aus der Graphik ersichtlich bilden diese Informationen die Grundlage für Szenarientwicklungen und Risikoanalysen. Ziel ist dabei immer eine Verbesserung der Handlungsfähigkeit der Gesamtorganisation.

---

<sup>4</sup> Vgl. Meurers, C., Göllner, J., Quirchmayr, G., Vogl, A.; „Wissensmanagement im ÖBH – Einführung in die Informationslogistik als Grundlage zur Wissens- und Organisationsentwicklung“, Schriftenreihe der Landesverteidigungsakademie, Band 20/2015, Wien, 2015, herausgegeben von Göllner, J., Meurers, C., Quirchmayr, G.

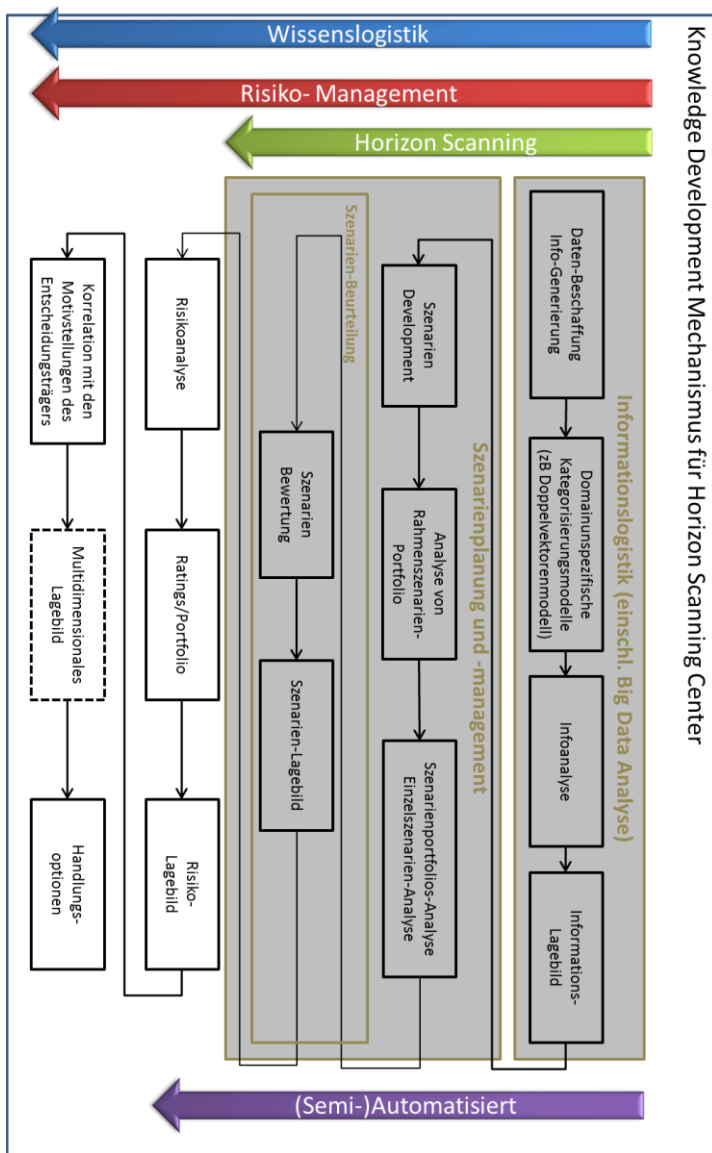


Abbildung 3: Wissensentwicklungsmechanismus: Z-Modell<sup>5</sup>

<sup>5</sup> Göllner, J., Klerx, J., Mak, K., Meurers, C.; „Wissensmanagement im ÖBH - Foresight in der strategischen Langfristplanung“, Schriftenreihe der Landesverteidigungsakademie, Band 5/2015, S.73, Wien, 2015, herausgegeben von Göllner, J., Klerx, J., Mak, K.

Von der Qualität des Wissenslogistik-Prozessclusters der Informationslogistik hängt ab, wie die Qualität der Szenario Analyse<sup>6</sup> und die Qualität der Risikoanalyse aussehen. Damit wird verdeutlicht, welche hohen Qualitätsanforderungen „Open Source Informationen“ gerecht werden müssen. In dieser Publikation wird im Folgenden dargestellt, wie die Wissenslogistik Prozesse des CDRC aufgebaut sind. Es wird dokumentiert, welche Innovationen es in den zwei Pilotphasen dazu gegeben hat und wie diese in die Prozesse der ZentDok integriert wurden.

Das langfristige Ziel des CDRC ist es einen Wissensentwicklungsmechanismus für ein kontinuierliches staatliches Horizon Scanning am Beispiel der Domain CYBER aufzubauen, entsprechend des Konzeptes staatlicher Horizon Scanning Center, dargestellt in der Publikation „Wissensmanagement im ÖBH, Foresight in der strategischen Langfristplanung“<sup>7</sup>

### **3 CDRC/ZentDok – Hauptaufgaben und korrespondierende Produkte**

Innerhalb der Produktlandkarte der ZentDok beteiligt sich das CDRC mit einer ganzen Reihe an wissensintensiven Produkten. Die folgende Darstellung gibt eine Übersicht über die Produkte des CDRC, die bereits vor Einrichtung des CDRC in der ZentDok in ähnlicher Weise hergestellt wurden.

---

und Second Design by Göllner, Klerx 07/2013 published in Klerx, J., Göllner, J., Mak, K., Horizon Scanning for emerging risks in supply chain systems, in: Wilby, Blachfellner, Hofkirchner (Eds.), Book of Abstracts, EMCSR-European Meetings on Cybernetics and Systems Research, S.601-607, Wien, 2014; vorgestellt im Rahmen der Präsentation auf der EMCSR 2014

<sup>6</sup> Göllner, J., Klerx, J., Mak, K., Meurers, C.; „Wissensmanagement im ÖBH - Foresight in der strategischen Langfristplanung“, Band 5/2015, Wien, 2015, S.75, herausgegeben von Göllner, J., Klerx, J., Mak, K.

<sup>7</sup> Göllner, J., Klerx, J., Mak, K., Meurers, C.; „Wissensmanagement im ÖBH - Foresight in der strategischen Langfristplanung“, Band 5/2015, Wien, 2015, S.80ff, herausgegeben von Göllner, J., Klerx, J., Mak, K.

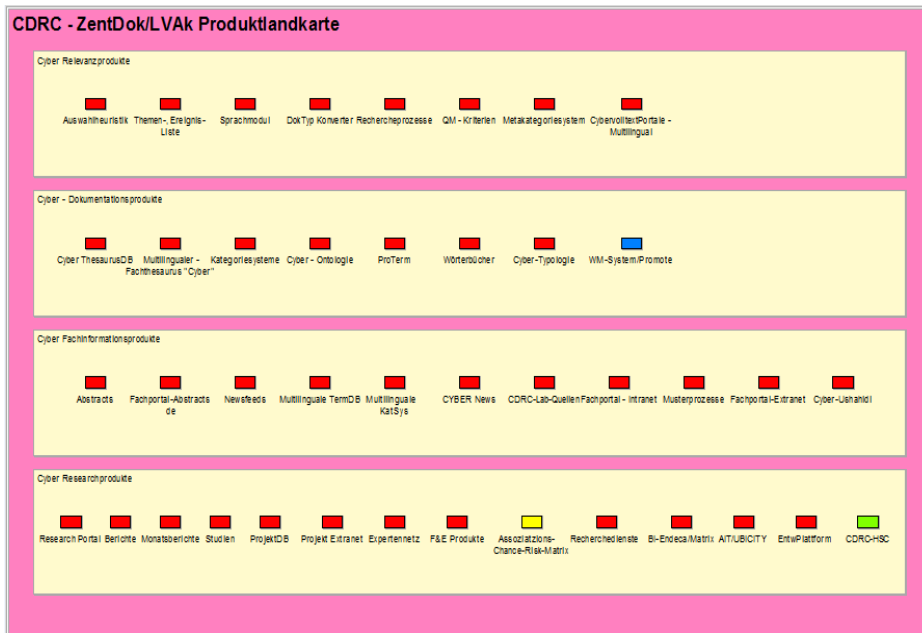


Abbildung 4: Produktlandkarte des CDRC in der ZentDok/LVak<sup>8</sup>

In weiterer Folge werden in Kapitel 3 jene vier Aufgabenbereiche und deren korespondierenden Produkte präsentiert, die neu sind und denen besondere Bedeutung in der Domain Cyber zu kommt. In der Produktlandkarte widerspiegeln sich die vier Aufgabenbereiche des CDRC. Zu den Hauptaufgaben des CDRC zählen dabei derzeit Beiträge zur:

- Identifizieren von OSInfo mit „Relevanz Monitoring“ mit Schwerpunkt Cyber
- Dokumentation und Translation von deutsch- und fremdsprachigem Texten mit Relevanz für die Cyber Domain
- Aufbereitung von Fachinformation
- Eigene Forschung sowie eigene Entwicklungen, die die Erfüllung der Hauptaufgaben verbessern oder effizienter gestalten.

<sup>8</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok



### 3.1 Identifizieren von OSInfo mit „Relevanz Monitoring“

Beim Relevanz Monitoring im CDRC geht es darum, möglichst schnell, zuverlässig und effizient Open Source Informationen (OSInfo) zum Thema „Cyber“ in Deutsch und in ausgewählten Fremdsprachen zu identifizieren und in einer Art zu verarbeiten, dass langfristig von einem kontinuierlichen Monitoring gesprochen werden kann.

Einige typische Problemstellungen dabei sind, dass das Internet eine solche Menge und Vielfalt an Informationen liefert, dass diese sowohl mit manuellen, als auch mit automatisierten Methoden nicht auswertbar sind. Für die manuelle Auswertung sind es zu viele Informationen und die automatisierte Auswertung scheitert an der semantischen Klassifikation und der hermeneutischen Interpretation.

Im CDRC wurde deswegen eine pragmatische Vorgehensweise gewählt, die sowohl die manuelle Auswertung, als auch die automatisierte Auswertung integriert, wie in der folgenden Darstellung zusammengefasst dargestellt wird. Dabei hat sich gezeigt, dass die Qualität durch eine anfängliche manuelle Quellenauswahl deutlich besser wird, als beim automatisierten Crawling, selbst wenn für das automatisierte Crawling „Smart Crawler“ eingesetzt werden.

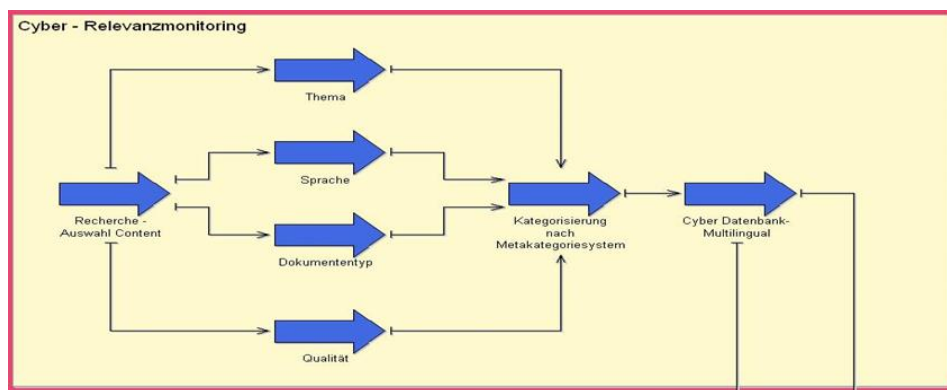


Abbildung 5: Prozessdiagramm Cyber-Relevanz Monitoring<sup>9</sup>

<sup>9</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok

Der Prozess besteht deswegen aus einer initialen Recherche zur Identifikation des relevanten Content mit iterativen Wiederholungsschleifen, bei denen sowohl neue Quellen dazu kommen können, als auch einmal identifizierte Quellen wieder aus dem Monitoring genommen werden können.

Die Ergebnisse des Monitorings werden nach folgendem Schema klassifiziert:

- Zuordnung der Inhalte zu Themen und Unterthemen (derzeit bereits 29 Kategorien!)
- Zuordnung zu Sprachen (derzeit 8 Sprachen getestet)
- Zuordnung zu Dokumententyp (Text, Bild, Multimedia, Social Media,...)
- Qualitätsbewertung...

In weiterer Folge dienen der manuell erarbeitete Content, die manuell identifizierten Quellen und die Klassifikation als Referenzdatensätze für ein automatisiertes Monitoring.

### **3.2 Dokumentation und Translation von mehrsprachigem Content**

Bei der Dokumentation und Translation von mehrsprachigem Content im CDRC geht es darum, die im Relevanz Monitoring entstandenen Informationen weiter zu verarbeiten. Je nach Sprache und Relevanzkategorie, wird diese Information einer automatisierten oder manuellen Übersetzung zugeführt.

Wie in der folgenden Abbildung ersichtlich, werden die Artikel nach der Übersetzung in mehrfacherweise ausgewertet. Nach der regelbasierten inhaltlichen Erschließung werden neue Begriffe zur Erweiterung der mehrsprachigen Fachthesauri verwendet. Diese dienen als Grundlage für eine zunehmende automatisierte Bearbeitung der entsprechenden Wissensdomain.

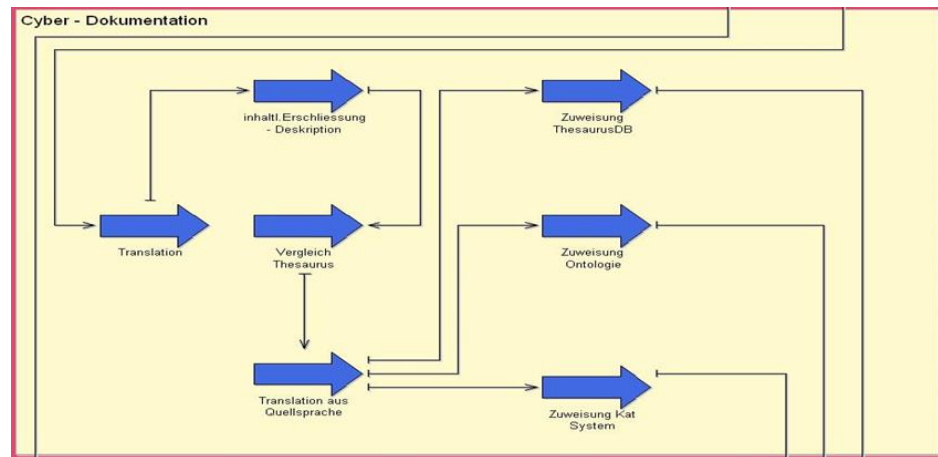


Abbildung 6: Prozessdiagramm Cyber-Dokumentation<sup>10</sup>

Im Einzelnen umfassen die Prozessschritte der Dokumentation und Translation folgende Tätigkeiten:

- Inhaltliche Erschließung (Deskription, regelbasiert)
- Aufbau von mehrsprachigen Terminologien zum Thema Cyber
- Pflege und Erweiterung der Thesauri, Taxonomien
- Aufbau und Erweiterung semantischer Netze, assoziativer Netze
- Aufbau und Verbesserung der Categoriesysteme
- Aufbau und Verbesserung von Ontologien
- Erstellung und Pflege von domainspezifischen Abkürzungsverzeichnissen

Alle oben angeführten Wissensprodukte dienen dazu, dass die semantische Erschließung einer Domain über die Zeit hinweg besser, schneller und zuverlässiger wird. Das aufgebaute Wissen kann dafür verwendet werden, um das Relevanz Monitoring zu verbessern, um die Datenbeschaffung zu automatisieren, um das Wissen darin besser zu indizieren und um das Wissen darin in weiterer Folge Anderen besser zugänglich zu machen.

<sup>10</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok

### 3.3 Fachinformation

In der Bereitstellung von Fachinformationen im CDRC geht es darum, das Wissen aus dem Relevanz Monitoring zu verdichten und zu interpretieren. Die Informationen aus unterschiedlichen Quellen werden damit in Bezug auf die darin enthaltene Information über Zusammenhänge ausgewertet. Speziell in der Cyber Domain zeigt sich hier der große Vorteil durch die Vorgehensweise beim CDRC. Im Cyber Bereich haben die Innovationen eine so hohe Geschwindigkeit, dass die Prozesse der klassischen wissenschaftlichen Kommunikation mit der Geschwindigkeit der Innovationen nicht mithalten können. Sowohl Bücher, als auch wissenschaftliche Paper sind zu langsam. Sogar Konferenzbeiträge hinken ab und zu hinter dem State of the Art hinterher. Internetquellen, mit Homepages, Blogs oder YouTube Videos enthalten deutlich neuere Informationen. Durch deren direkte Auswertung und Bereitstellung über die Zusammenfassungen zu Fachinformationen ist die Aktualität der Information sicher gestellt.

Wie aus dem Prozessdiagramm ersichtlich werden die Informationen aus der Cyber Volltextdatenbank verwendet um Abstracts zu erstellen, die in die Abstract DB einfließen. Weiters werden die Informationen verwendet, um Folgerecherchen zu starten und um einen Newsfeed zu befüllen. Alle verdichteten Informationen sind über Cyber Portale dem spezifizierten Nutzerkreis des CDRC zugänglich.

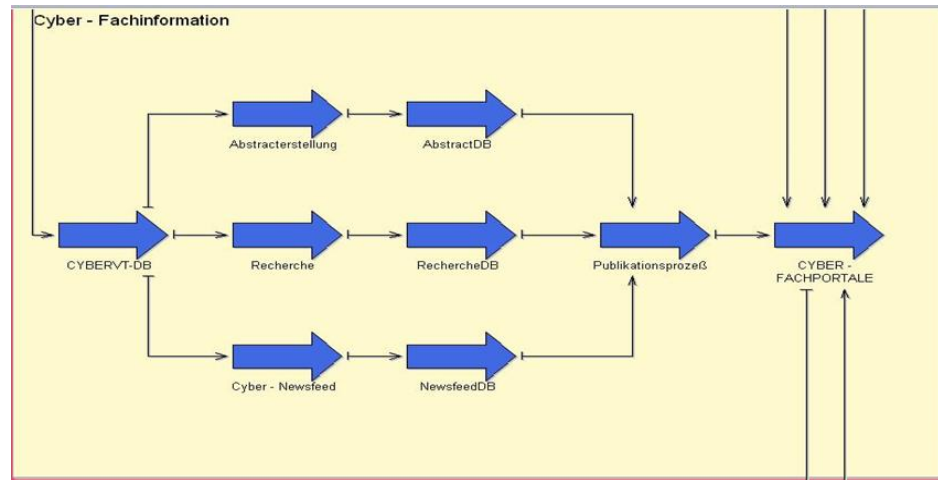


Abbildung 7: Prozessdiagramm Cyber-Fachinformation<sup>11</sup>

Die Hauptaufgabe „Bereitstellung von Fachinformationen“ umfasst im CDRC folgende Tätigkeiten:

- Aufbau von Referenz- & Volltextdatenbanken zum Thema Cyber
- Erstellung von deutschsprachigen Abstracts spezieller Dokumente
- CROWD-Recherchedienste in Referenzdatenbank
- CROWD-Recherchedienste in mehrsprachiger Volltextdatenbank
- Zurverfügungstellung von Newsfeeds
- Aufbau und Update von Intranet und Extranet Informationsplattformen
- Aufbau eines „Cyber News-Dienstes“

Mit der Bereitstellung der Fachinformationen wird es den Usern des CDRC im Laufe der Zeit sukzessive immer besser ermöglicht einen tagesaktuellen Stand des „State of the Art“ in verschiedensten Cyberthemen abzufragen. Damit wird eine Grundlage geschaffen, dass bei Forschung und Entwicklung sowohl bestehendes aktuelles Wissen

<sup>11</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok

verfügbar ist, als auch Wissen über „Unwissen“ daraus abgeleitet werden kann.

Dieses Wissen über die Lücken der Forschung, über Forschungsfronten, Emerging Technologies und über aktuelle Innovationen ermöglicht erst, gezielt relevante neue Forschung zu betreiben. Auch wenn in der Wissenschaft immer wieder die Relevanz unterschiedlicher Sprachen leichtfertig abgetan wird, so ist es doch ein nicht wegzudiskutierendes wertvolles Ergebnis aus dem Pilotbetrieb des CDRC, dass die wissenschaftlichen Erkenntnisse insbesondere aus dem asiatischen, aber auch aus dem russischen, spanischen und dem arabischen Sprachraum für den Cyber Bereich auch im Westen relevant sind. Bisher wurden diese Informationen über die klassischen wissenschaftlichen Kommunikationskanäle viel zu langsam oder gar nicht kommuniziert. Mit der multilingualen Fachinformation bietet sich eine Möglichkeit, dieses zu verbessern.

### **3.4 Forschung und Entwicklung**

Die vierte und letzte Hauptaufgabe des CDRC ist es, in Kooperation mit externen Partnern Forschung und Entwicklung zu betreiben. Nachdem alle Produkte des CDRC im Kern wissensintensive Services sind, ist eigene Forschung und Entwicklung unumgänglich. Eine Kooperation mit Universitäten und Forschungsorganisationen der angewandten Forschung bieten ein Umfeld, in dem neue Innovationen entstehen und direkt umgesetzt werden können.

Wie das Prozessdiagramm zeigt, werden jeweils vom Forschungsmanagement Forschungsprojekte, Studien und Berichte initiiert, die zum nachhaltigen Wissensaufbau des CDRC beitragen. Die Ergebnisse aus diesen Projekten werden dokumentiert und fließen in die Forschungsdatenbank des CDRC ein.

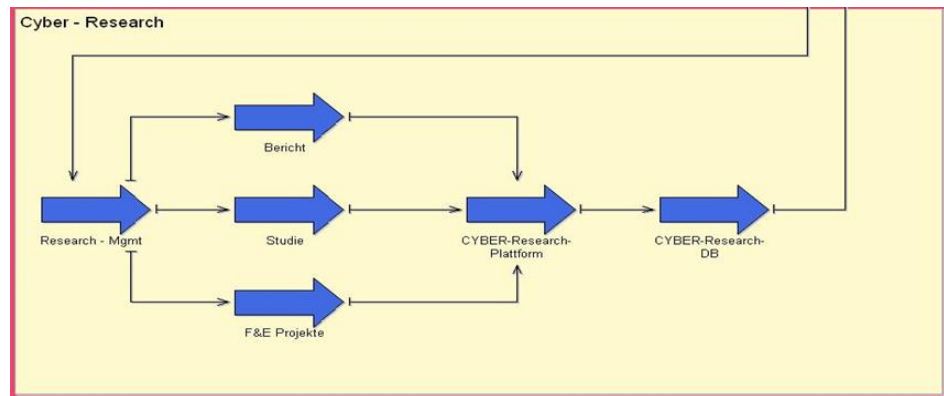


Abbildung 8: Prozessdiagramm Cyber-Research<sup>12</sup>

Aus dieser Tätigkeit ergeben sich folgende forschungsrelevante Produkte die im Kontext des CDRC erstellt werden:

- Erstellen von „Cyber Faktendokumenten“
- Erstellen von Studien, Diplomarbeiten und Dissertationen
- Mitarbeit bei und Leitung von F&E Projekten
- Betreuung von wissenschaftlichen Arbeiten
- Aufbau und Update eines „Cyber Research Portals“
- Aufbau eines „Cyber Informations-Extranets“
- Aufbau und Pflege eines Expertennetzes,
- Aufbau „Expertenstab Cyber“, „Cyber Miliz“
- Erstellung von Cyber-Berichten (6W)
- Erstellung einer Assoziations-Matrix (Chance-Risiko-Matrix)
- Erstellung von Monatsberichten
- Experten Tagging Horizon Scanning Center „Cyber“
- Entwicklung von OSInfo-relevanten Anwendungen
- Mitarbeit bei Entwicklung von Werkzeugen und Methoden sowie Services des „CDRC-ZentDok/LVAK“

Der Betrieb von eigener Forschung und Entwicklung spiegelt die konsequente Umsetzung des Konzeptes zum Aufbau einer effektiven

<sup>12</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok

Wissenslogistik gemäß dem eingangs erwähnten Z-Model wieder. Nachdem Wissen in der Cyber Domain die entscheidende strategische Ressource ist, spiegelt sich in den Hauptaufgaben des CDRC der effiziente Umgang mit dieser Ressource wieder. Die Fachinformationen tragen dazu bei, den State of the Art und die Forschungsfronten zu dokumentieren. Daraus ergeben sich Hinweise auf offene Forschungsfragen, sowohl in direkter als auch in indirekter Weise.

In direkter Weise ist Forschung und Entwicklung immer dann notwendig, wenn die entsprechenden Tools, Methoden und Konzepte für die Tätigkeit des CDRC noch nicht ausreichend vorhanden sind. Der daraus entstehende Bedarf für Eigenentwicklungen wird soweit möglich durch interne Kapazitäten, aber auch durch externe Kooperationen abgedeckt.

In indirekter Weise entstehen Hinweise und Ideen für neue Forschungsaktivitäten aus der Diskrepanz zwischen State of the Art und den existierenden Lösungen im gesamten Cyber Bereich. Nachdem die Aufgaben des CDRC allerdings eng an die Zielsetzung und die damit verbundenen Kernkompetenzen gebunden sind, werden diese Hinweise auf mögliche und sinnvolle Innovationen im Cyber Bereich den Nutzern in eigener Verantwortung zur Verfügung gestellt, mit der Absicht, dass diese vom dem spezifischen Wissen ihren eigenen Nutzen realisieren können.

In den folgenden zwei Kapiteln werden Innovationen des CDRC präsentiert, die aus der konsequenten Umsetzung der eigenen Forschung und Entwicklung entstanden sind, und die die Effizienz des CDRC maßgeblich verbessert haben.

#### **4 CDRC Innovation: Crowd OSInfo mit CROWD**

Die Auswertung von Open Source Quellen mit Hilfe von Gruppen ist an sich nicht neu. Allerdings wurde diese Auswertung bisher nicht mit einer explizit formulierten Informationslogistik unterlegt, die dazu entwickelt worden ist, durch enge Kooperation in der Zeit die Gruppenleistung sukzessive zu verbessern.

Sowohl staatliche Dienste, als auch Wissenschaftler und Techniker arbeiten üblicherweise mit Experten, die durch ihre jahrelange Ausbildung und



Erfahrung für die Qualität ihrer Arbeitsleistung garantieren. Der Erfolg dieses Prozesses ist davon abhängig, dass der einzelne Experte in der jeweiligen Situation seine Leistung erbringt. Eine Qualitätssicherung ist nur über eine Art Peer Review mittels Urteil von anderen Experten möglich.

In den letzten Jahren hat es vermehrt Erfolge damit gegeben, wissensintensive Services entweder mittels Computern automatisiert, oder über Computer Interfaces menschliche Urteilsfähigkeit so einzubinden, dass diese den Teil der Wissensarbeit übernehmen, in denen Computer notorisch schlecht sind (Semantische Interpretation, Hermeneutische Interpretation, Urteilsfindung und Begründung, inhaltliche Bilderkennung, Klassifikation nach Thema). Plattformen, wie Mechanical Turk von Amazon, Clickworker, Cloudcrowd, CastingWords, blur Group, GeniusRocket, Whinot, IdeaOffer, NamingForce, SquadHelp, Threadless, 99designs, Crowdspring, Prova, CrowdTap und viele andere bieten Ansätze das Leistungspotential menschlicher Gruppen im Kontext von Crowd Working<sup>13</sup> abzurufen.

Viele dieser Crowd Working Ansätze sind geeignet, um die Schwierigkeiten des Computers bei der inhaltlichen Auswertung von Texten zu verbessern. Deswegen werden diese Ansätze in traditioneller Ausprägung auch schon lange beim Informationsmanagement staatlicher Services eingesetzt. China nutzt z.B. die Erkenntnisse seiner Bürger, die sich im Ausland aufhalten<sup>14</sup>. Im arabischen Kontext werden moderne Formen der Informationsbeschaffung<sup>15,16</sup> und Verbreitung durch Swarmcast<sup>17</sup> auch im Kontext ziviler Gruppen umgesetzt, die sich im Internet bewegen. In den USA gibt es Hinweise, dass Ideen zu HUMINT und OSINT mittels Crowd zumindest existieren<sup>18</sup>. Mit dem Ausblick dieser Publikation, vom Mai

---

<sup>13</sup> Vgl. URL: <http://mashable.com/2011/04/01/make-money-crowdworking/>, [23.09.2015]

<sup>14</sup> Vgl. Mattis, P., L.; "Li Kenong and the Practice of Chinese Intelligence", in: *International Journal of Intelligence and CounterIntelligence*, Volume 28, Issue 3, S.540-556, 2015

<sup>15</sup> Vgl. Westerman, T.; "TERROR GOES ANALYTICAL - AL QAEDA IS WATCHING, Same Hatred, New Approach, looking for Weakness, and No Apologies", in: *International News Analysis Today*, 2004, URL: <http://inatoday.com/terror%20goes%20analytical%20032604.htm>, [8.9.2015]

<sup>16</sup> Vgl. Sanchez, S., E.; "Spider Web: Al-Qaeda's Link to the Intelligence Agencies of the Major Powers", in: *International Journal of Intelligence and CounterIntelligence*, Volume 28, Issue 3, S.429-448, 2015

<sup>17</sup> Vgl. URL: <http://www.thedailybeast.com/articles/2014/10/01/isis-is-winning-the-online-jihad-against-the-west.html>, [23.09.2015]

<sup>18</sup> Vgl. Stottlemire, S., A.; "HUMINT, OSINT, or Something New? Defining Crowdsourced Intelligence", in: *International Journal of Intelligence and CounterIntelligence*, Volume 28, Issue 3, S.578-589, DOI: 10.1080/08850607.2015.992760

2015: „*I suggest that KDM crowdsourcing, separate from Daren Brabham's other types, should be professionalized across sectors ...*“<sup>19</sup>, fordert Steven A. Stottlemyre, genau das, was im CDRC vollzogen und hier in der Publikation schon dokumentiert wird.

Das CROWD OSInfo im CDRC ist ein spezifischer neuer Ansatz zum Informationsmanagement und stellt eine Innovation des CDRC dar, welche im Folgenden im Detail vorgestellt und auf ihre Wirksamkeit hin diskutiert wird.

Die Crowd des CDRC besteht gegenwärtig aus jeweils Gruppen von 10 bis 15 Personen mit größtenteils technisch sehr gut ausgebildeten HTL-Absolventen der Fachrichtungen Informations-, Nachrichten- oder Elektrotechnik. Ausnahmen im Pilotprojekt waren auch Absolventen der Technischen Universität Wien und von Gymnasien oder Handelsakademien die aber für die Gruppe eine Bereicherung darstellten und auch weiterhin vermehrt die „technische“ Crowd verstärken werden.

Zweites Hauptaugenmerk bei der Auswahl der Gruppen wurde auf sprachliche Fähigkeiten gelegt, die im ersten Jahr englisch, russisch, spanisch, chinesisch, indisch, polnisch, ukrainisch, serbisch und deutsch umfassten. In weiterer Folge ist ein Ausbau der Sprachkompetenzen vorgesehen und mit arabisch, ungarisch, slowenisch und persisch bereits geplant worden.

Letztendlich muss man von einer „hochqualifizierten Crowd“ sprechen, die dem herkömmlichen Vorstellungen von Crowds mit hohen Teilnehmerzahlen gegenübergestellt wurde. Der Begriff „High Quality - oder High Performance Crowd“ könnte hier entsprechend verwendet werden.

Weiters beeinflusste der kulturelle Hintergrund der Gruppen maßgeblich den Wert des Projektes. Nach mehreren Recherchen zu unterschiedlichen Fragestellungen kam es zur systematischen Abbildung der Zugänge der Gruppe zu verschiedenen Themenbereichen. Dies widerspiegelt sich im nachfolgenden Bild der Assoziationsheuristik der CDRC-CROWD.

---

<sup>19</sup> Stottlemyre, S., A.; „HUMINT, OSINT, or Something New? Defining Crowdsourced Intelligence“, in: International Journal of Intelligence and CounterIntelligence, Volume 28, Issue 3, S.587, DOI: 10.1080/08850607.2015.992760

#### **4.1 CROWD – Assoziationsheuristik**

Die Crowd Assoziationsheuristik bildet die Grundlage für das Verständnis der Funktionsweise der Crowd. Es hat sich gezeigt, dass je nach kultureller, sprachlicher oder personeller Zusammensetzung die OSInfo Ergebnisse der Crowd sowohl qualitativ, als auch quantitativ ganz unterschiedliche sind. Um diese Unterschiede zu verifizieren und zu erklären hat sich die Dokumentation in Form einer Assoziationsheuristik bewährt.

Vor allem die Faktoren persönliche Kompetenz, Sprache, Vorwissen, Interessen, Persönlichkeitsmerkmale und Zeit spielen eine bestimmende Rolle zur Erklärung der Ergebnisse einer Crowd Recherche. Die folgende Graphik gibt einen detaillierten Überblick über die Merkmale, die bestimmend für die die Ergebnisse der Crowd Recherchen waren.

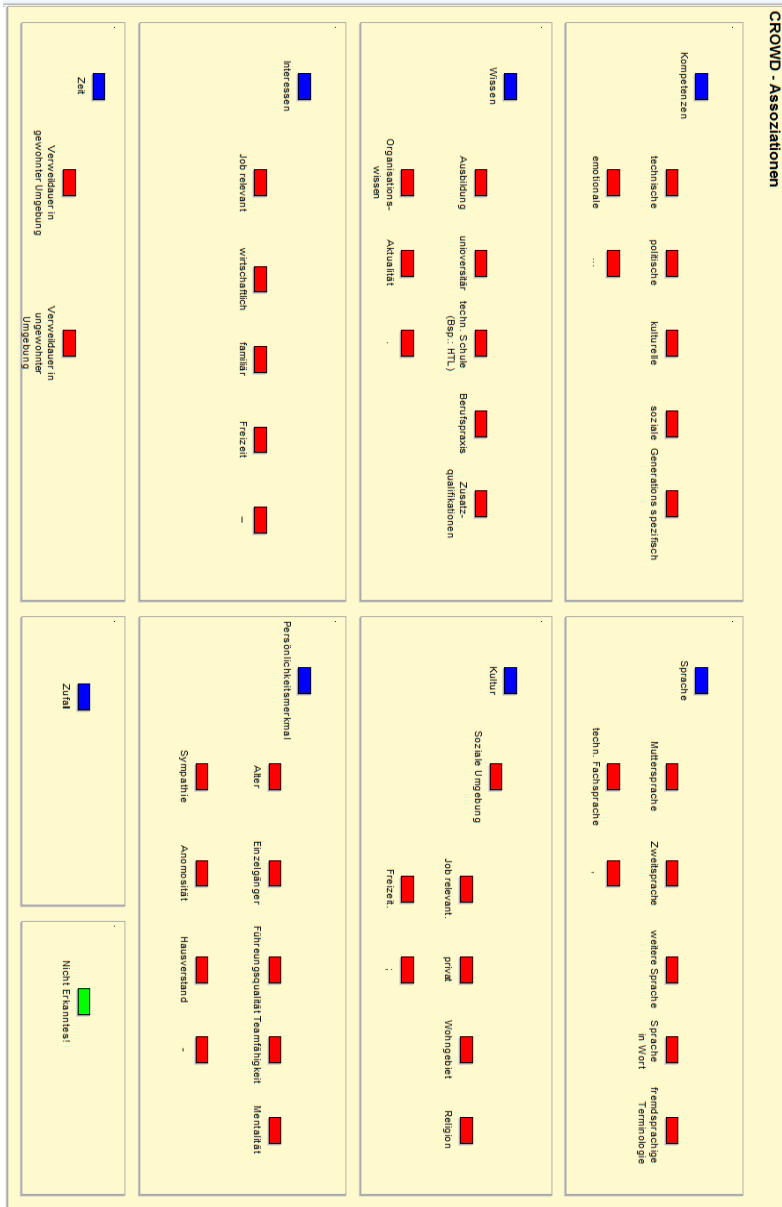


Abbildung 9: Übersicht Crowd Assoziationsheuristik<sup>20</sup>

<sup>20</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok

Alle dargestellten Assoziationen traten mehr oder weniger ausgeprägt bei der Beantwortung der Fragestellungen in bestimmten Konfigurationen auf. Diese Dokumentation kann dynamisch in PROMOTE® weiterentwickelt werden und zur Qualitätssicherung wesentlich beitragen. Nach jeder Recherche ist es so möglich zu überprüfen, ob neue Sichtweisen zu einer Fragestellung eingeflossen sind oder ob eine Erweiterung der Recherche unter bisher nicht berücksichtigten Assoziationen, die bereits in der Matrix dokumentiert wurden, notwendig erscheint. Diese Qualität konnte bisher auch durch Experten in keiner Weise auch nur annähernd erreicht werden.

Bei Analyse der Fragestellungen wurde ebenso systematisch die Vielfalt erfasst und abgebildet. Es wurden mehr als dreihundert Fragen analysiert, die in den letzten Monaten aus allen Nutzerkreisen an die ZentDok gestellt wurden. Ebenso wurden spezielle Anforderungen der Fragen in der Domain Cyber in dieser Heuristik mitdokumentiert.

Auf der Grundlage dieser Ergebnisse lässt sich die Zusammensetzung der Crowd optimieren. Je nach Fragestellung, Themengebiet, oder Zielsetzung der Recherche und nach Rahmenbedingungen, wie Zeit, Ressourcen, usw., hat sich eine spezifische Zusammensetzung bewährt. Ziel dabei sollte es sein, mit möglichst wenig Aufwand eine hinreichend vollständige Recherche zu produzieren. Zur Beurteilung der Qualität der Recherche wurden sowohl Referenzrecherchen, als auch Beurteilungen von Experten eingesetzt.

#### **4.2 ZentDok – Fragenheuristik**

Ein weiteres zentrales Merkmal zur Optimierung einer Crowd Recherche ist die Fragenheuristik. Die Art der Fragestellung bestimmt ganz entscheidend den Verlauf und die Ergebnisse von Crowd Recherchen. Zudem bieten sie einen Überblick über Interessen und Motivationen der User. Die Auswertung der Fragenheuristik erlaubt es mit der Zeit die Qualität der Crowd Recherchen deutlich zu verbessern.

Fragenklassen lassen sich z.B. nach Art der Frage, (Fragen zu Problemstellungen, Zusammenhängen, bestehenden Sachverhalten, Ursachen, Wirkungen, usw.) klassifizieren. Daraus ergeben sich offene und breite Recherchen mit viel Interpretationsspielraum für die Crowd.

Fragestellungen zur Beschreibung und Erklärung von Ereignissen (wer, wann, was, wo, wie, warum), treten häufig im journalistischen Kontext auf. Sie regen zu spezifischen Recherchen an, die zu einer strukturierten Erklärung eines Ereignisses führen.

Fragestellungen die sich auf einen Zeitverlauf beziehen, wie zeitliche Entwicklung, mögliche Entwicklungen (Zukunft), vorangegangene Entwicklungen und Erklärungen, periodische Regelmäßigkeiten, Trends, usw. haben sich als besonders schwierig erwiesen. Vor allem die Suche nach zukünftigen Entwicklungen bedarf spezieller Suchstrategien. Die Crowd ist deutlich effizienter, wenn diese Suchstrategien im Vorhinein vermittelt wurden.

Themenorientierte Fragestellungen, wie Krisen, Konflikte, Chancen und Innovationen, Erkenntnisse, usw. haben eine ähnliche Dynamik, wie ereignisbezogene Suchen, sind jedoch weniger strukturiert. Je nach Thema ist der Verlauf ganz unterschiedlich. Bei den meisten Themenklassen ist eine Verbesserung durch Übung zu erreichen.

Die Qualität bei Crowd Recherchen zu technische Fragestellungen, wie z.B. State of the Art, Patente, Statistiken, Anwendungsbeispiele, Vergleiche und Funktionsweisen, hängt von Vorbildung und Interessen des Crowd Recherche Teams ab. Technisch vorgebildete Teammitglieder verwenden ganz andere Suchstrategien als nicht technisch Vorgebildete. Es hat sich jedoch gezeigt, dass sprachlich oder anderweitig vorgebildete Teammitglieder die Ergebnisse durch ihre unkonventionellen Suchstrategien sehr bereichern können.

Fragestellungen zu Organisation, Management und Strategien, wie Führungsverhalten, Pläne, Roadmaps, Führungsverfahren usw. ergeben häufig eine Ergebnisliste, die weit umfangreicher ist, als in der Fragestellung nahegelegt. Bei einer Qualitätssicherung durch Experten reduzieren sich diese Ergebnisse.

Fragestellungen zur Forschungsrelevanz, mit Literaturrecherchen, Programmausschreibungen, Projektberichten, Forschungsinstituten, Expertennetzwerken, MOOCs, Entwicklungsplattformen, usw. erfordern spezifische jedoch auch unkonventionelle Quellen. So hat sich gezeigt, dass

zunehmend auf Videoplattformen, wie z.B. YouTube, auch relevante Ergebnisse zu diesem Segment zu finden sind.

Insgesamt führt die Pflege von Frage-Heuristiken dazu, dass Fragen spezifischer an die Crowd gestellt und formuliert werden können. Bei Einführungen in die Internetrecherche hat sich bewährt, dass durch Suchheuristiken die Bandbreite der Suchmethoden gut erläutert werden kann. Durch die Möglichkeit des Vergleiches der unterschiedlichen Sichtweisen auf Sachverhalte mit der Vielfalt der Fragestellungen ergaben sich bisher kaum systematisch erfasste Kombinationsmöglichkeiten, die sich auch dynamisch weiterentwickeln. In dieser Form stellen sie in Kombination mit Unterstützungsmechanismen, wie bspw. Terminologieframeworks, Retrievalfunktionalitäten oder Textanalysetools, eine neue Herausforderung für die Analysekapazitäten der Nutzer dar. Durch weitere Forschungs- und Entwicklungsaufgaben sollen diese Mechanismen besser erforscht und verbesserte Verarbeitungsmöglichkeiten im OSInfo Bereich geschaffen werden. Dies stellt ein weiteres Schwergewicht in der zukünftigen Arbeit der ZentDok gemeinsam mit ihren Partnern dar.

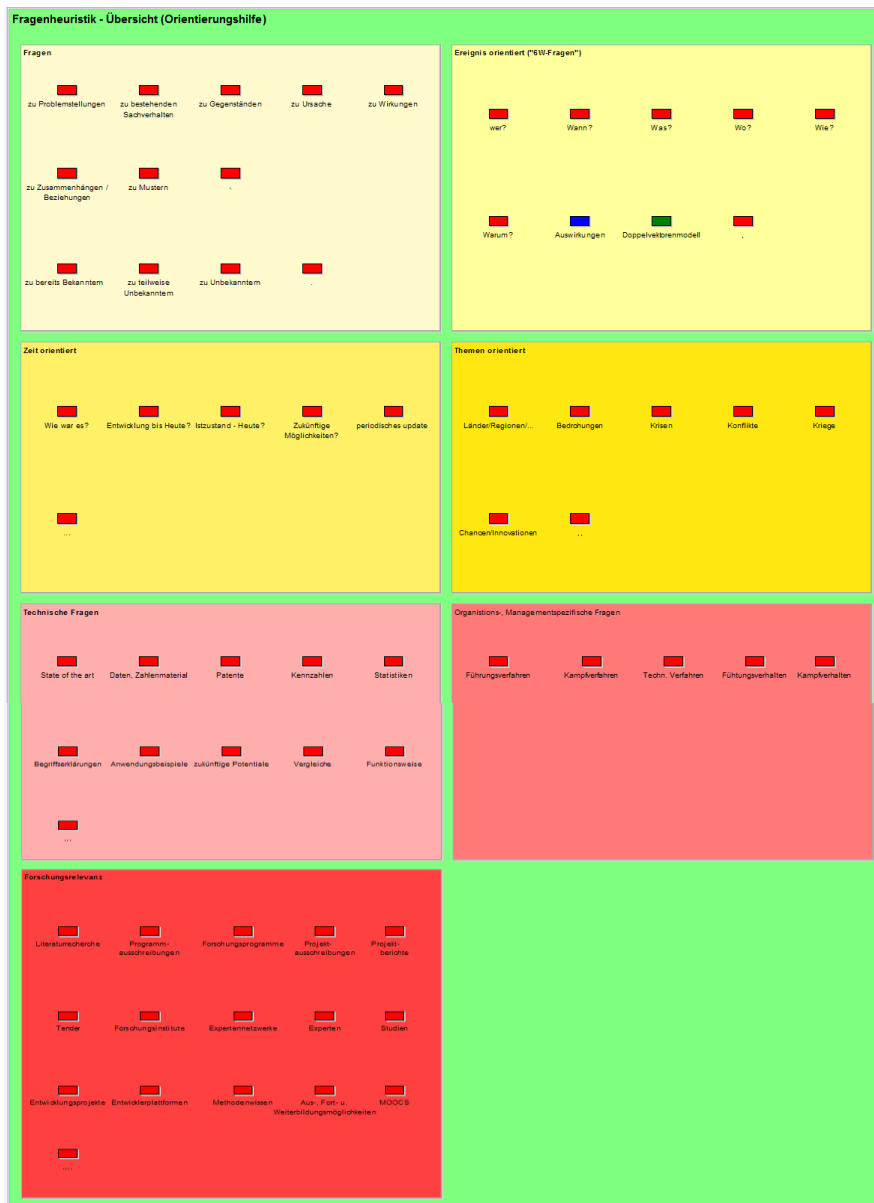


Abbildung 10: Übersicht Crowd Suche Fragenheuristic<sup>21</sup>

<sup>21</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok



### 4.3 CROWD – Recherche

Assoziationsheuristik und Fragenheuristik bieten die kontrollierte Grundlage für den Einsatz eines Teams zur Crowd Recherche. In diesem Kapitel wird der Ablauf einer Crowd Recherche im Detail vorgestellt und Vor- bzw. Nachteile diskutiert. In der wissenschaftlichen Literatur konnten, abgesehen von den Publikationen und Software Entwicklungen zu Crowd Working<sup>22</sup> und Crowd Intelligence<sup>23</sup>, keine Hinweise auf den Einsatz einer organisierten Crowd zur Recherche gefunden werden.

Die Möglichkeit eine hoch „performante“ Gruppe („High Quality Crowd“ oder „High Performance Crowd“) auf eine Fragestellung zu fokussieren, ergab einige überraschende Ergebnisse. Durch die hohe Qualität der Ausbildung, durch die vielfältigen Sprachkenntnisse und die in der Matrix abgebildeten weitläufigen Assoziationen, waren die Ergebnisse nicht nur quantitativ sehr überzeugend, sondern in der Qualität und Vielfalt so anspruchsvoll, dass sie auch bei Expertenrecherchen bei weitem nicht ähnliche Ergebnisse erreicht werden konnten. Auch der Zeitfaktor wurde zu einem wesentlichen Kriterium, der durch diese Gruppen bereits um einen Faktor 15 bis 20 schneller Ergebnisse lieferte, als ein einzelner Forscher bei Vergleichstests erreichte.

Dabei wurde bisher deutlich, dass zwar mehr und vielfältigere Ergebnisse gefunden wurden, es konnte jedoch noch nicht geklärt werden, wo die Grenzen dieser Verbesserung liegen. Die Abdeckung einer Fragestellung, um zu „vollständigen“ Ergebnissen zu kommen, wird ein weiterer wesentlicher Forschungsgegenstand der Zukunft.

Die CROWD Recherchen werden unterstützt von Musterprozessen, die in der ZentDok bereits abgebildet waren und während des Projektes weiterentwickelt wurden. Die Dokumentation mit Prozessdiagrammen hat sich nachhaltig bewährt und ersparte massiv an Ausbildungszeit. Im Folgenden wird der im CDRC entwickelte Crowd Recherche Musterprozess vorgestellt und Erfahrungen aus der Anwendung diskutiert.

---

<sup>22</sup> Vgl. URL: <http://mashable.com/2011/04/01/make-money-crowdworking/>, [23.09.2015]

<sup>23</sup> Vgl. Stottlemire, S.,A.; „HUMINT, OSINT, or Something New? Defining Crowdsourced Intelligence“, in: *International Journal of Intelligence and CounterIntelligence*, Volume 28, Issue 3, S.578-589, DOI: 10.1080/08850607.2015.992760

### **4.3.1 CROWD Recherche - Musterprozess**

Der Crowd Recherche Musterprozess fasst die zentralen Prozessschritte der Crowd Recherche aus der Perspektive eines Team Mitgliedes der Crowd zusammen. Er dient als Grundlage für die Schulung und als Vorlage für die Vorgehensweise bei einzelnen Recherchen. Ergänzt wird dieser Musterprozess durch das Wissen über die Crowd Heuristik und die Fragen Heuristik.

Die nachfolgende Graphik bietet einen Überblick über die einzelnen Prozessschritte.

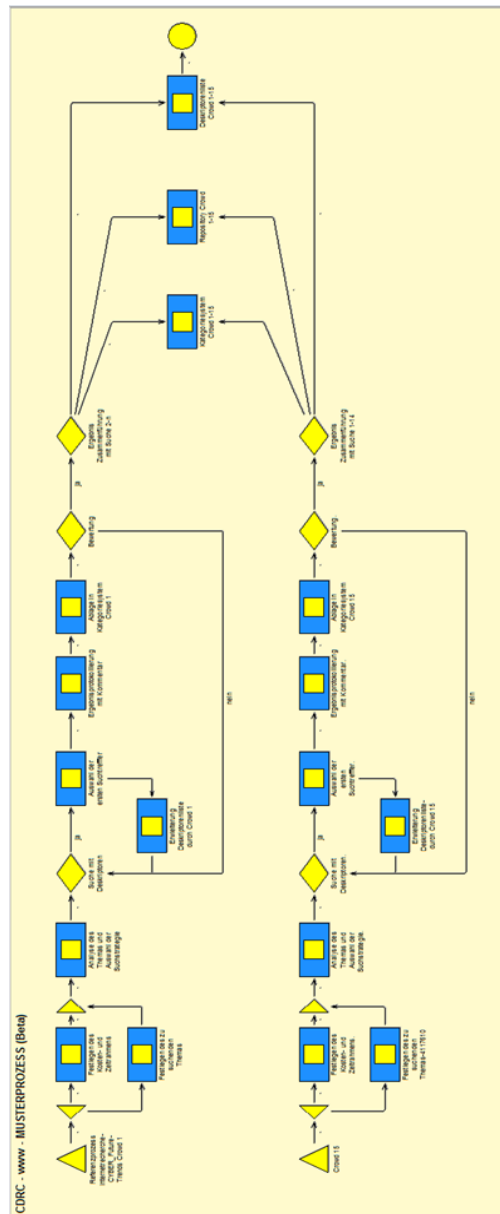


Abbildung 11: Prozessdiagramm Crowd Recherche<sup>24</sup>

<sup>24</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok

Die Prozessschritte einer Crowd Recherche bestehen aus der Festlegung der Suchzeit, der Präsentation der vorbereitenden Analyse sowie der Strukturierung des Themas durch den Experten an die Crowd. Diese wählen während der Präsentation des Experten ihre Suchstrategie aus und notieren sich diese mit den korrespondierenden Suchbegriffen. Daraufhin erfolgt die Recherche Auftragserteilung an die Crowd.

Die Crowd dokumentiert ihre Ergebnisse laufend, während der Recherche mit und kategorisiert die Ergebnisse gemäß dem in Kapitel 5.2.1 dokumentierten Kategorisierungsschema (Beispiel für Cyber). Nach einer Zusammenfassung, Bewertung und Kontrolle durch Fachexperten, erfolgt die formale Meldung aller Einzelrecherchen mit dem CDRC Meldesystem.

Es hat sich gezeigt, dass eine zweite Recherche nach der Präsentation der Suchergebnisse und der Suchstrategien vom Experten zur Crowd die Ergebnisse deutlich verbessern. Dieser zweite Iterationsprozess der Crowd Recherche läuft, abgesehen von den zusätzlichen Ausgangsinformationen, wie der erste Prozess ab.

Bei Festlegung von Zeitrahmen und Suchthema korreliert die notwendige Zeit mit der Offenheit der Themenspezifikation und dem Typus der Fragestellung. Je präziser beides ist, umso kürzer kann der notwendige Zeitrahmen ausfallen. Allerdings führt eine präzise Themenspezifikation auch zu weniger Variationen in den Ergebnissen. Je nach Fragestellung kann dies ein Vorteil oder ein Nachteil sein. Die folgende Graphik zeigt den ersten Teil der Crowd Recherche, bis zur Zusammenführung der Ergebnisse.

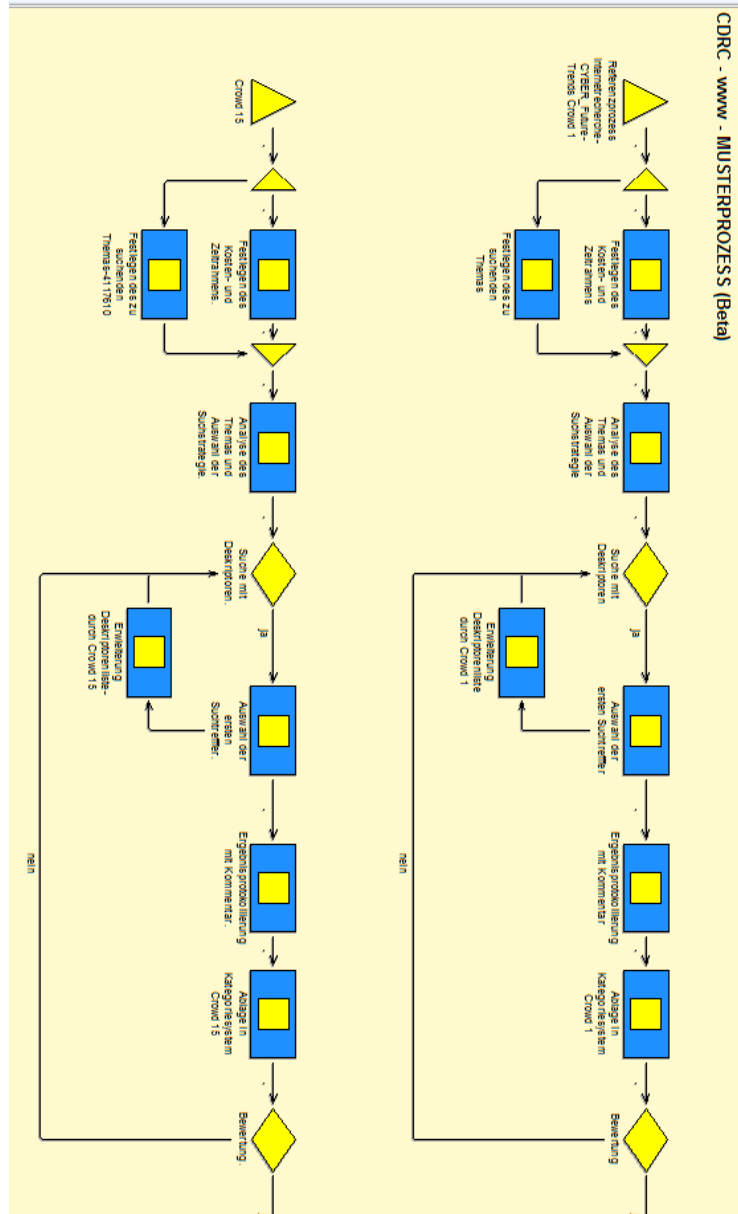


Abbildung 12: Prozessdiagramm Crowd Recherche Teil 1: Crowd<sup>25</sup>

<sup>25</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok

Im ersten Teil des Crowd Recherche Prozess Diagramms wird gezeigt, wie der Recherche Auftrag an die Crowd erfolgt und wie die Suche durchgeführt wird. Wesentlich für die Qualität ist dabei die verständliche und strukturierte Formulierung des Rechercheauftrages für „nicht Fachexperten“. Es hat sich bewährt, die Präsentation der vorbereiteten Analyse und Strukturierung des Themas durch den Fachexperten dafür zu verwenden, dass Thema verständlich und strukturiert an die Crowd zu übermitteln. Damit es nicht zu Missverständnissen kommt, ist es wichtig, der Crowd in dieser Phase so viel wie möglich auch Hintergrundinformation über Anlass und Zweck der Recherche zu übermitteln. Die Auswahl und Bestimmung der Suchstrategie wird daraufhin eigenständig von der Crowd vorgenommen. Eine Nahelegung spezifischer Suchstrategien durch den Experten würde die Assoziationsfähigkeit der Crowd begrenzen.

Daraufhin erfolgt die Auftragserteilung an die Crowd mit einem zeitlichen Limit. Dieses sollte nicht zu lange sein, da nach einer gewissen Zeit das Austesten von neuen Suchstrategien zunehmend langsamer wird. In dieser Phase hat sich bewährt, dass der Fachexperte anhand der ersten Ergebnisse überprüft, ob es bei der Darstellung des Themas nicht zu Missverständnissen gekommen ist und bei Bedarf den Rechercheauftrag nachjustiert. Die Recherche Ergebnisse werden von der Crowd mit Protokollierung und Kategorisierung der Ergebnisse entsprechend dem CDRC Kategorisierungsschema (Kapitel 5.2.1) in das CDRC Meldesystem hochgeladen und stehen dort für eine weitere Verwendung zur Verfügung.

Es hat sich gezeigt, dass eine zweite Recherche nach der Präsentation der Suchergebnisse und der Suchstrategien vom Experten zur Crowd die Ergebnisse deutlich verbessern. Dieser zweite Iterationsprozess der Crowd Recherche läuft, abgesehen von den zusätzlichen Ausgangsinformationen, wie der erste Prozess ab.

Aus der Crowd Recherche heraus entstehen neben den Ergebnissen der Recherche auch weitere wertvolle Wissensprodukte, wie z.B. ein domainspezifisches Kategoriensystem und domainspezifische Deskriptoren Listen. Diese müssen von Zeit zu Zeit zusammengeführt und adaptiert werden. Im folgenden Teil 2 des Prozessdiagramms wird dieser Prozess dargestellt.

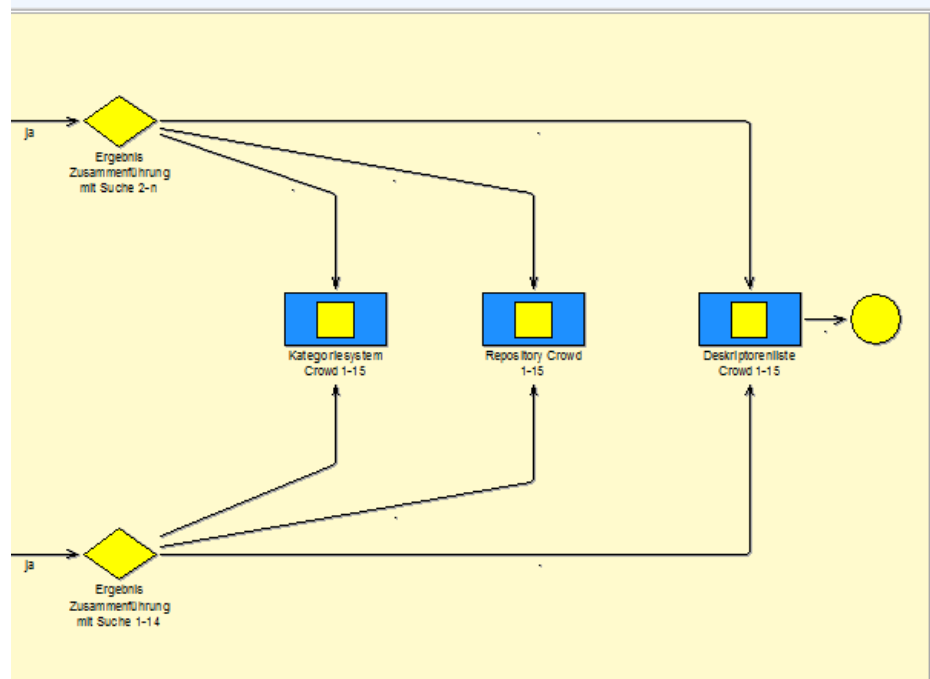


Abbildung 13: Prozessdiagramm Crowd Recherche Teil 2: Zusammenführung<sup>26</sup>

Da die Zusammenführung der Ergebnisse auch einen Überblick über die erreichte Qualität bietet, ist dieser ein Teil des Qualitätsmanagements. Das Ziel des Qualitätsmanagements im Crowd OSInfo ist, die Relevanz, die Richtigkeit der Recherche Ergebnisse sicher zu stellen. Dazu müssen Informationen von drei Urteilebenen zusammen fließen. Zum einen beurteilen die Mitglieder der Crowd bei jedem Suchergebnis, ob dieses relevant, richtig oder sogar wertvoll für eine weiterführende Verarbeitung ist. Dann beurteilen die Experten im CDRC die Ergebnisse der Crowd Recherche. Und schließlich beurteilt der Auftraggeber noch einmal diese Ergebnisse. Alle Urteile enthalten wertvolle Informationen für die Qualitätssicherung und sollten deswegen dokumentiert und gespeichert werden.

Zusätzlich konnten schon eine ganze Reihe an Internet Services identifiziert werden, die direkt bei der Recherche ein schnelles verifizieren

<sup>26</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok

erlauben. Es gibt z.B. verschiedenen Möglichkeiten nicht mehr direkt im Web verfügbare Informationen doch noch zu beschaffen, z.B. mit Suchmaschinen Cache oder über die Wayback Engine<sup>27</sup>. Einzelne Services können bei Detailrecherchen und verifizierenden Recherchen helfen, wie z.B. das citizen evidence lab<sup>28</sup>, welches Anleitungen bietet, wie Internetmeldungen verifiziert werden können und einen Service anbietet um Metainformationen über YouTube Videos zu bekommen.

Aus der Suche nach den „wertvollsten Informationen“ bieten High Profile Sites eine besondere Rolle. HPS haben sich bei wiederholten Recherchen immer wieder aus verschiedenen Gründen als zuverlässige und wertvolle Informationsquelle gezeigt. Deswegen werden sie im CDRC nicht nur in der Qualitätssicherung bei Crowd Recherchen verwendet. Sie unterliegen auch einer besonderen Behandlung, sowohl bei der Identifizierung, als auch bei der Verarbeitung. Im folgenden Kapitel wird dieses beschrieben.

### **4.3.2 High Profile Sites (HPS)**

Wesentliches Kernelement im Qualitätsmanagement stellen die sogenannten „High Profile Sites“ („HPS“) dar. Diese Sites beinhalten wesentliche Informationen zur Domain Cyber, die regelmäßig durch die Cyber-Rekruten ausgewertet werden. Beurteilt werden diese Sites durch die jeweiligen internen und externen Experten nach den unten abgebildeten Kriterien, die in dieser Form erstmals systematisch erfasst wurden.

---

<sup>27</sup> Vgl. URL: <http://archive.org/web/>, [23.09.2015]

<sup>28</sup> Vgl. URL: <http://citizenevidence.org/>, [23.09.2015]



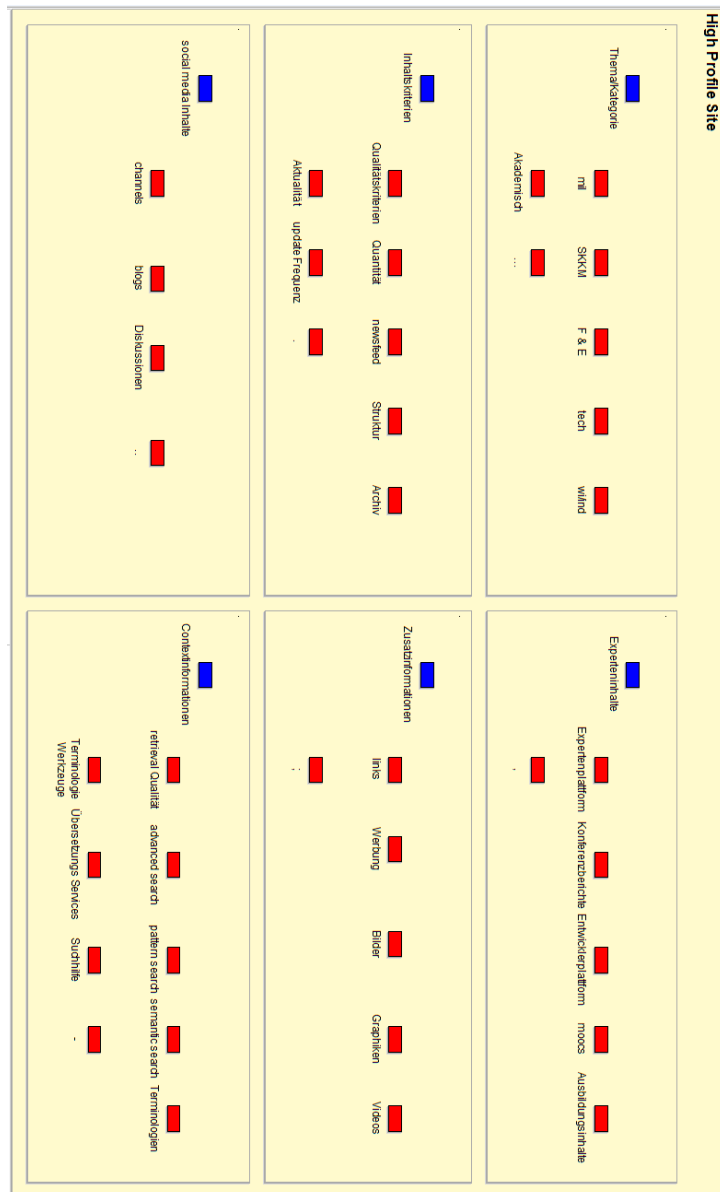


Abbildung 14: Kriterienkatalog von High Profile Sites<sup>29</sup>

<sup>29</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok

HPS werden bei normalen Recherchen identifiziert, oder in spezifischen Recherchen extra gesucht und identifiziert. Sollten diese bei anderen Tätigkeiten im CDRC auftauchen, so sind alle Mitarbeiter des CDRC angehalten diese zu melden und damit einer speziellen Bearbeitung zuzuführen.

Grundsätzlich zeichnen sich HPS dadurch aus, dass sie thematisch relevante Informationen für die zentralen Kategorien des CDRC enthalten. Weiters werden diese Informationen immer wieder aktualisiert. Häufig haben sich diese Sites wegen ihrer Qualität als zentraler Hub für eine Community etabliert.

HPS mit aktuellen Experteninhalten sind z.B. in der Community etablierte Expertenplattformen mit Konferenzberichten, die regelmäßig aktualisiert werden, mit Entwicklerplattformen, wo Software und Tools ausgetauscht und regelmäßig aktualisiert werden.

HPS aufgrund von Inhaltskriterien (Qualität, Quantität, Aktualität, update Frequenz) sind z.B. Social Media Plattformen, die immer wieder kehrend neue und aktuelle Informationen hervor bringen, wie z.B. YouTube Channels, Expertentwitter, oder Expertenforen.

HPS aufgrund von Kontextinformationen über Suchmaschinen, Recherchertools, Terminologiewerkzeuge und Übersetzungsservices sind z.B. Sites über investigativen Journalismus, Hactivismus oder andere zivilgesellschaftliche Aktivitäten aus dem Cyber Bereich.

Die HPS dienen zum kontinuierlichen Wissensaufbau im CDRC und werden in Crowd Recherchen nicht nur als inhaltliche Quelle, sondern auch für aktuelle methodische Anregungen verwendet.

#### **4.4 Cyber Berichte des CDRC (6W)**

Die Cyber Berichte des CDRC sind eine Form der Zusammenfassung der Ergebnisse der Crowd Recherchen, die sich an der Eventbezogenen Auswertung orientiert. Durch die hohe Qualität der Rekruten konnten nach relativ kurzer Einweisung sogenannte 6W-Berichte erstellt werden. Dabei werden Informationen zu bestimmten Ereignissen über Hintergrund-Internet-Recherchen gesammelt und von zwei unabhängigen

Recherche-Teams ausgewertet. Danach erfolgt eine Zusammenfassung (W6) und ein Vergleich beider Berichte, die wiederum diskutiert und zu einem endgültigen Bericht konsolidiert werden.

Die sechs W-Fragen stellen sich dar wie folgt:

1. Wer? Wer sind die Akteure? Wer ist betroffen?
2. Was? Was ist passiert?
3. Wann? Zu welchem Zeitpunkt? Welcher Zeitraum?
4. Wie? Wie fand das Ereignis statt? Wie kam es dazu?
5. Wo? Wo geschah es? Welche Länder sind betroffen?
6. Warum? Welche Beweggründe gab es? Welche Motive?

Die Ergebnisse fließen neben der Lieferung an die jeweiligen Auftraggeber, wie alle anderen Inhalte auch, in ein sog. „Cyber Repository“, wo alle Inhalte, Metadaten, Terminologien und Services für weitere Verarbeitungsschritte zur Verfügung gestellt werden.

#### **4.5 CROWD – Assoziations-Matrix-Prozess („Chance – Risiko“)**

Die Crowd Assoziationsmatrix ist ein weiteres Produkt zur Auswertung der Crowd Recherchen. Bei der Recherchetätigkeit entsteht ein intrinsisches Wissen zur möglichen Chancen und Risiken im Cyber Bereich. Daher ist es naheliegend zu versuchen, dieses Wissen in seiner manifesten Form und mit geeigneten Methoden von der Crowd abzufragen und sichtbar zu machen. Die Möglichkeit, Entwicklungs- und Forschungsaufgaben durch zu führen, ergab verschiedenste Möglichkeiten, die dazu getestet werden konnten. Eine besonders interessante Perspektive auf zukünftige Potentiale für den Einsatz von mehreren Crowds zeigt uns der Assoziations-Matrix-Prozess.

Die Gruppe bewertete kategorisierte Inhaltsbereiche (wie bspw. Internet of Things, ...) nach der Möglichkeit darin Risiken oder Chancen (Potentiale) für Sicherheit, Innovationen oder wirtschaftlichen Nutzen zu finden, auf einer Skala von eins bis zehn. Zuvor wurden in den jeweiligen Hauptkategorien (derzeit 29!) auch Unterkategorien bestimmt, die normalisiert wurden.

Im dargestellten Prozess werden die einzelnen Schritte dokumentiert und sind mit Detailinformationen zu den jeweiligen methodischen Erfordernissen ausgestattet. So wird beispielsweise der Auswertelgorithmus im Detail festgehalten. Die nachfolgende Graphik gibt einen Überblick über diesen Prozess.

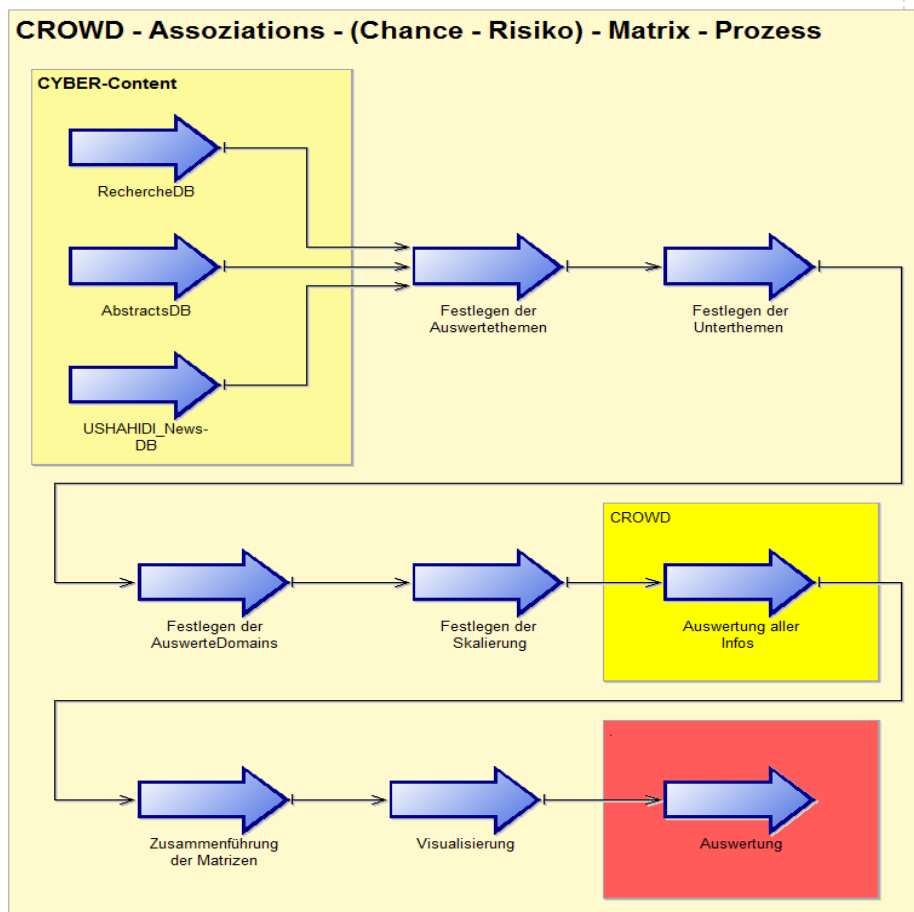


Abbildung 15: CROWD – Assoziations-Matrix-Prozess („Chance – Risiko“)<sup>30</sup>

Nach einer Objektivierung der Ergebnisse kommt es zu einer Visualisierung mit einer Testumgebung, wo insbesondere Hauptaugenmerk

<sup>30</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok

auf Themen und Unterthemen zu legen war, die nicht in ein typisches Muster passten.

Alle Ergebnisse dieser Testanwendungen gaben Anlass, dass Crowd Assoziationsmuster und die Ergebnisse genauestens auf die Folgen und weitere Untersuchungsgegenstände hin zu analysieren sind.

Durch Einzelpersonen erscheint es nicht annähernd möglich zu sein, komplexe Themenbereiche und Sichtweisen in dieser Form auszuwerten und Neues darin sichtbar werden zu lassen – egal ob Chancen oder Risiken. Auch die Möglichkeit, Ursache-Wirkungsmechanismen besser nachvollziehen zu können, muss als wesentlicher Vorteil dieser Methode beachtet werden.

#### **4.6 Visualisierung der CROWD – Assoziations-Matrix**

Die Crowd Assoziations-Matrix ermöglicht es prinzipiell eine Bewertung von Chancen und Risiken zu Themen auf beliebigem Abstraktionsniveau von einer großen Gruppe von Menschen einzusammeln. Ziel dabei ist es nicht die Crowd für komplette SWOT Analysen einzusetzen und schon gar nicht ein Risikomanagement der Experten zu ersetzen. Vielmehr ist die Crowd Assoziations-Matrix eine Methode um Chancen Risiko Einschätzungen einer Crowd abzufragen und einer weiteren Bearbeitung hinzuzuführen.

in diese Form ist die Visualisierung ein Ausgangspunkt für weitere Entwicklungsschritte und Testung von neuesten Methoden und Verfahren um Anomalien und andere sich verändernde Muster zu erkennen, die in bisherigen Analysemodellen nicht erkennbar waren. Die Ergebnisse liegen auf unterschiedlichen thematischen Abstraktionsleveln vor und können auch auf allen Abstraktionsleveln visualisiert werden. Die folgende Graphik zeigt die Chancen Risiko Bewertung für Cyber auf dem obersten Abstraktionslevel.



Abbildung 16: Chancen Risiko Matrix im Bereich Cyber<sup>31</sup>

<sup>31</sup> Quelle: eigene Darstellung

Aus der Graphik ist zu erkennen, dass es einen schwach positiven Zusammenhang zwischen Chancen und Risiken gibt. Auf der hohen Abstraktionsebene sollte das nicht überbewertet werden. Aber es zeigt, dass Themen wie Big Data, Militär und Politik mehr aus der Perspektive der Chance wahrgenommen werden. Themen wie Terrorismus, Big Data und Cyber Security dagegen eher aus der Perspektive des Risikos. In den Unterthemen hat sich gezeigt, dass dieser Zusammenhang sich zum Teil auflöst. Das kann als Hinweis verstanden werden, dass das Labeling immer auch etwas über die Erwartung hinsichtlich der Chancen und Risiken transportiert.

Big Data nimmt eine Sonderposition ein, da es sowohl mit Chancen als auch mit Risiken wahrgenommen wird (Chancen aus den Erfolgen der Auswertung von hoch faktoriellen Datensätzen und Risiken aus der Überwachungsperspektive). Dieses Beispiel macht die Problematik von SWOT Analysen im Kontext des CDRC deutlich. Chancen für die Einen, können Risiken für die Anderen darstellen. Daraus ergibt sich ein Forschungsbedarf in Bezug auf Crowd OSInfo und Risikomanagement.

## **5 Zusammenführung von ZentDok-Experten- & CROWD-Rechercheprozess**

Es hat sich in der Pilotphase des CDRC gezeigt, dass die Ergebnisse von Crowd OSInfo Analysen wie folgt integriert werden können. Oben angesprochene Ergebnisse führten zu folgender Zusammenführung der Experten Recherche und der CROWD-Recherche bis zu einem Weiterverarbeitungsschritt, bei dem alle Inhalte und Strukturen für eine Analyse zur Verfügung stehen.

Damit ermöglicht eine vorangestellte Frageanalyse und die Nutzung aller Fähigkeiten sowie eine umfassende Qualitätskontrolle auch eine Weiterentwicklung der Muster und der terminologischen Unterstützungswerkzeuge. Diese sind dynamisch im Gesamtsystem anpassbar.

Der integrierte Prozess der parallelen Recherche (Expert/Crowd) startet dabei mit einer Fragestellung und einer Analyse der Fragestellung. Dann trennen sich die Prozesse. Der Experte greift auf seine Datenbanken

zurück und zieht bei Bedarf das Internet zu Rate. Die Crowd konzentriert sich mit dem Hintergrund der Zusatzausbildung auf die Internetrecherche unter Verwendung der dort gebotenen Recherche Tools. Beide werden durch Recherche Tools, wie Cyber Thesaurus, Cyber Ontologie und Übersetzungswerkzeuge unterstützt.

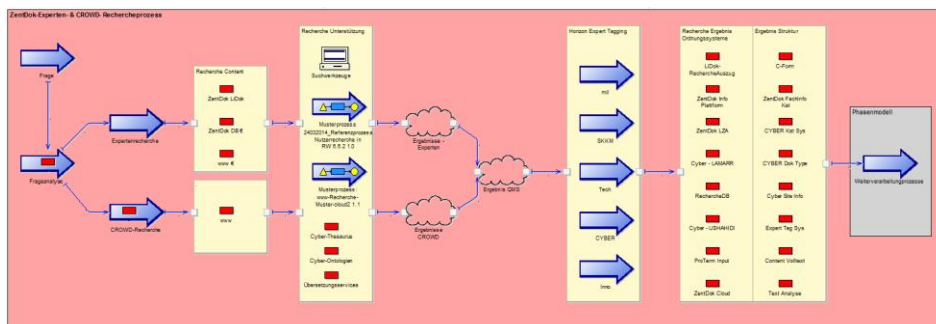


Abbildung 17: Prozessdiagramm Integration Experteneinschätzung/ Crowd Recherche<sup>32</sup>

Durch die Kombination beider Recherche Ergebnisse entsteht, unter der Verantwortung des Experten, eine qualitätsgesicherte Recherche, die verschiedene Standpunkte, Motivationen, Suchstrategien und Sprachen mit berücksichtigt. Die qualitätsgesicherte Recherche wird bei Bedarf vom Experten direkt für seine Analyse weiter verwertet. Gleichzeitig fließt die Recherche nach einem Expert Tagging in die CDRC Informationsplattform ein, wo sie nach verschiedenen Ordnungssystemen für spätere Auswertungen gespeichert wird. Der Schritt des „Horizon Expert Tagging“ sorgt dafür, dass die in der jeweiligen Domain speziell für ihre Nutzer angepasste „Tags“ zur Anreicherung des Informationsgehaltes und zur Verbesserung der Qualität der Ergebnisse gesetzt werden. Tags können z.B. auf die im Foresight Bereich relevanten weak signals oder auf andere spezielle Wissens Elemente hindeuten.

In der folgenden Abbildung wird der erste Teil des Prozesses der Zusammenführung von Expert Recherche und Crowd Recherche vorgestellt.

<sup>32</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok



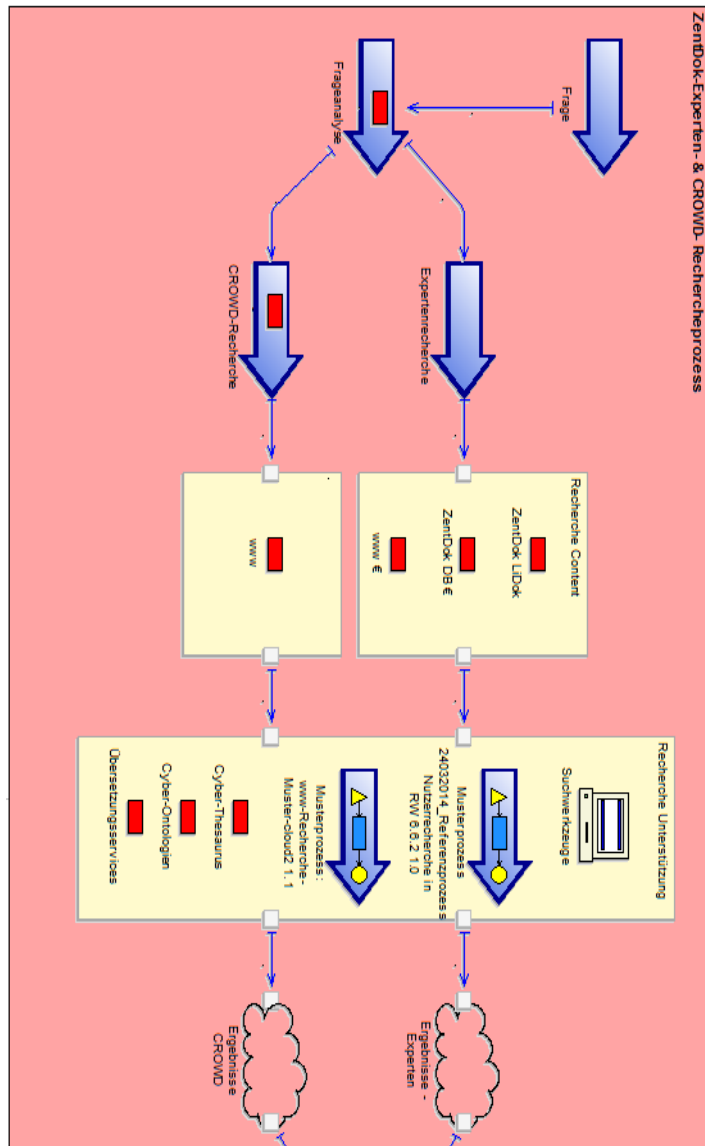


Abbildung 18: Prozessdiagramm zur Zusammenführung Detail A<sup>33</sup>

<sup>33</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok

Daraus ist ersichtlich, dass die Experten der ZentDok gemäß Musterprozess zur Datenbankrecherche auf Datenbanken wie LiDok und ZentDok DB zurückgreifen. Die Crowd Recherche ist im Dokument www-Recherche-Muster-cloud2 1.1 beschrieben. Auf den Cyber Thesaurus, die Cyber Ontologien und Übersetzungsservices können beide zugreifen.

Im Prozessdiagramm Detail B wird die Zusammenführung der Recherchen, über Horizon Expert Tagging bis hin zur Befüllung der Ordnungssysteme der ZentDok (inklusive Cyber LAMARR, Cyber Ushahidi, ProTerm und ZentDok Cloud) sowie die Strukturierung der Ergebnisse beschrieben.

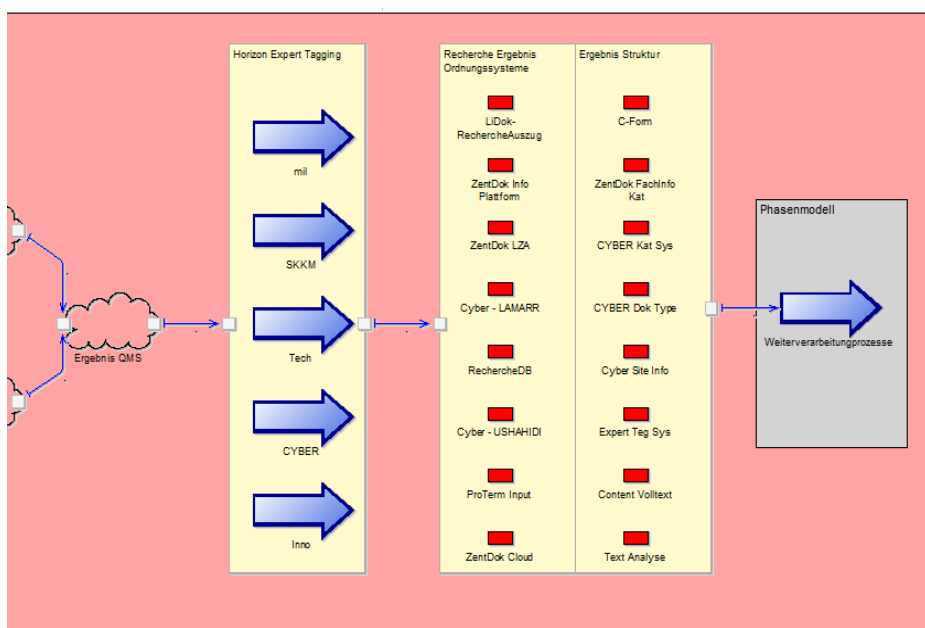


Abbildung 19: Prozessdiagramm zur Zusammenführung Detail B<sup>34</sup>

Damit stehen in der Ergebnis Verarbeitung qualitätsgesicherte Daten zur Verfügung die über Textanalyse, oder Volltext Indizierung gesucht und verwendet werden können.

<sup>34</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok

Die Weiterverarbeitung erfolgt prinzipiell anhand des Eingangs präsentierten Phasenmodells zur Wissensentwicklung, im Z-Modell. Nach der Zielsetzung des CDRC hat jedoch die Wissensbeschaffung und die Dokumentation derzeit eine höhere Priorität. Ohne eine qualitativ hochwertige Beschaffung sind weder Szenarien Analysen noch Risikoanalysen sinnvoll einzusetzen.

### **5.1 CDRC Innovation: Modell der zeitlichen Quelleninterdependenz zur Informationsverdichtung**

Inhaltlich steht hinter dem Prozess der Zusammenführung der Rechercheergebnisse des CDRC ein Modell der zeitlichen Quelleninterdependenz zur qualitätsgesicherten Informationsverdichtung (MQqI) im OSInfo Bereich, welches am CDRC entwickelt wurde. Durch die Herausforderung, offene Informationen für verschiedenste Nutzer zeitgerecht in hoher Qualität zur Verfügung zu stellen war es im CDRC notwendig die OSInfo Analyse auf den folgenden drei Hauptebenen zu starten:

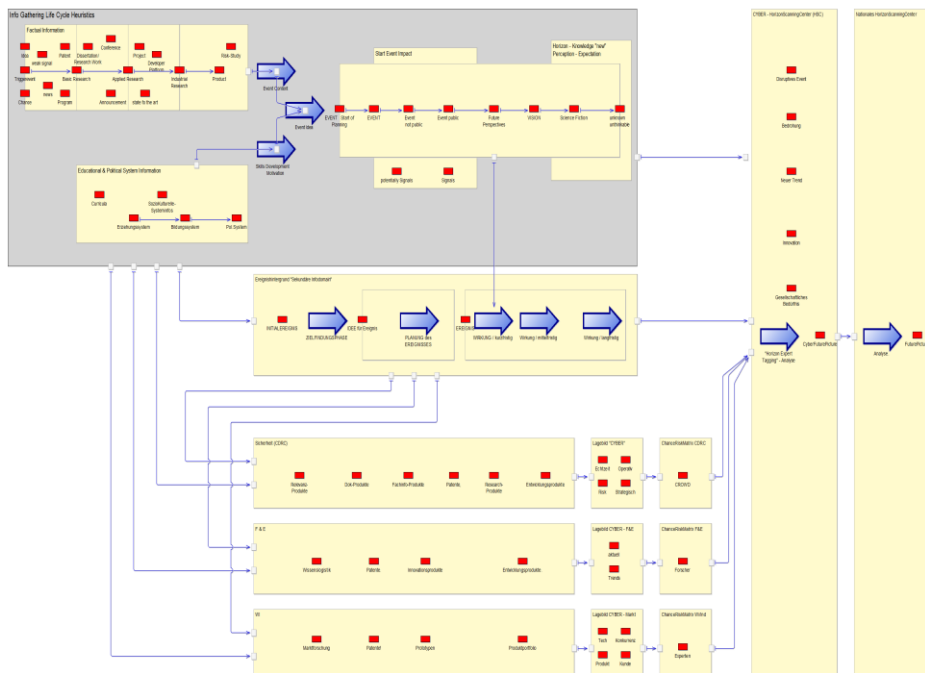
1. die Ebene der Domain Sicherheit (mil, civic),
2. die Ebene der Forschung und Innovation
3. sowie die Ebene der Technik (Cyber State of the Art).

Alle Inhalte der OSInfo Analysen dienen additiv zur Generierung von unterschiedlichsten Lagebildern. So werden Echtzeitlagebilder wie auch operative Lagebilder aber auch strategische Lagebilder oder Foresight Lagebilder mit unterschiedlichsten offenen Informationen angereichert oder diese darin verarbeitet.

Bei der Festlegung der Dokumententypen („DocType“) bei Crowd OSInfo Recherchen war aufgefallen, dass verschiedene Arten von Information quantitativ und qualitativ in einer Zeitschiene angeordnet werden können, um daraus zeitlich konsistente Entwicklungen abzuleiten.

Diese Erfahrung in der Crowd OSInfo Analyse führte zur Entwicklung des in der Abbildung verdeutlichten Modells der zeitlichen Quelleninterdependenz. Wenn die Informationen aus unterschiedlichen Quellen zu

unterschiedlichen Zeiten zur Erklärung von Ereignissen beitragen können, dann kann dieses Wissen als Grundlage für ein besseres und vor allem auch frühzeitiges Verständnis von neuen Entwicklungen beitragen.



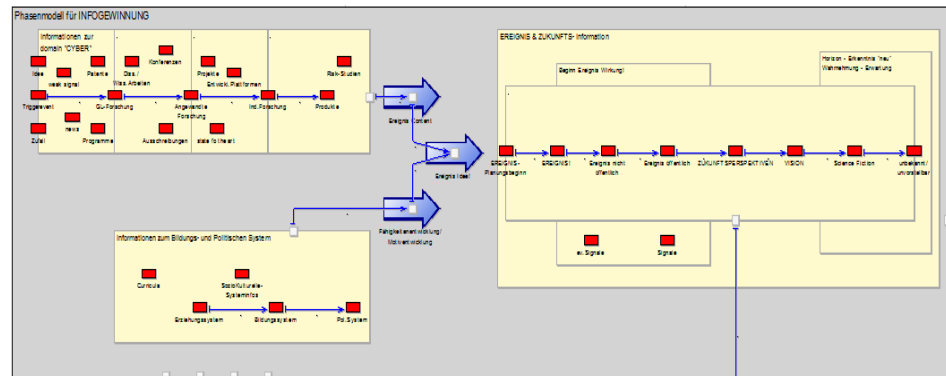
**Abbildung 20: Modell der zeitlichen Quelleninterdependenz zur Informationsverdichtung<sup>35</sup>**

Um dieses Wissen nutzen zu können ist es notwendig, den relevanten Wissensraum in einer ganz speziellen Art zu strukturieren. Nimmt man ein Ereignis als Ausgangspunkt für die zeitliche Zuordnung von Inhalten, so kann man bereits bei Zuordnung der jeweiligen „DocTypes“ während des Relevanz Monitorings die Geschichte zum Ereignis nachvollziehen. Diese Sachinformationen der Domain Cyber bspw. bilden auch die Möglichkeit ab, Entwicklungsgeschwindigkeiten zu analysieren. Wie lange braucht ein Patent zu einem Produkt? Oder wie lange braucht ein Produkt um bei einem relevanten Ereignis verwendet werden zu können? In der Domain Cyber spielt dieser Zeitfaktor eine wesentliche Rolle.

<sup>35</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok

Deswegen wurden zunächst im Phasenmodell der Infogewinnung alle im CDRC aufgetretenen Wissensobjekte so aufgetragen, dass diese in einem groben zeitlichen Zusammenhang stehen. Das Segment „Bildung/Politische Systeme“ aus der Infogewinnung hat dabei einen Sonderstatus und wurde heraus genommen, genauso, wie das Segment der Ereignisse und Zukunftsinformation.

Wie aus der Graphik ersichtlich tragen alle Wissensobjekte auf ihre spezifische Art und zu einem bestimmten Zeitpunkt zum Ablauf eines Ereignisses bei. Im Zentrum stehen dabei Ereignis Kontext, Ereignis Idee und die Fähigkeiten Entwicklung.



**Abbildung 21: Modell der zeitlichen Quelleninterdependenz  
Detail A: Infogewinnung<sup>36</sup>**

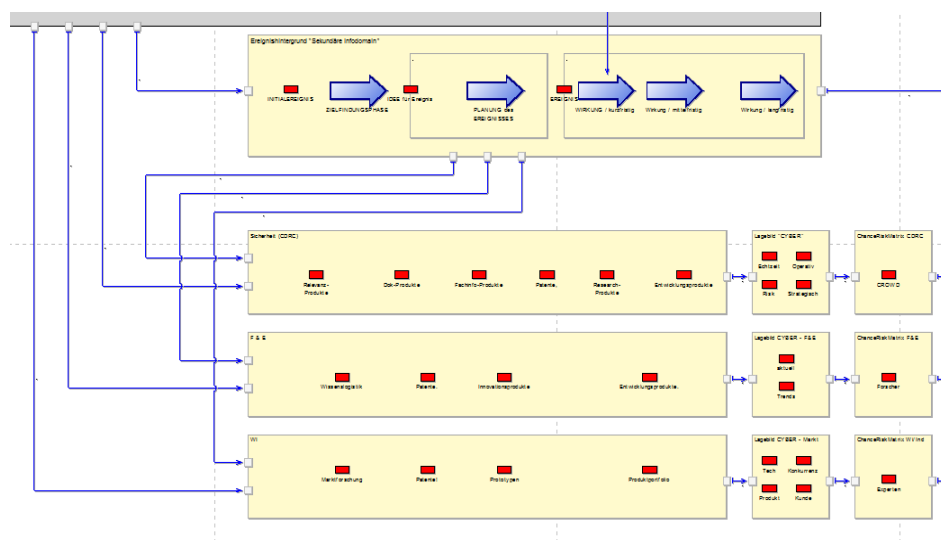
Im mittleren Bereich des Modells erkennt man diese als Wirkungsebene von Ereignissen. Drumherum findet sich die Darstellung einer umfassenden Sichtweise auf ein Ereignis als Diagramm für ein Gesamtbild rund um ein Schlüsselereignis und umliegende mögliche Abhängigkeiten.

Die Wirkung der Ereignisse, unmittelbar, kurz- und langfristig, die möglichen Signale, das erste öffentliche Auftreten, dies sind alles wesentliche Bausteine für die zukünftigen Potentiale – Chancen aber auch Risiken. Zukunftsstudien, Science Fiction Vorstellungen und Ähnliches unterstützen den Blick in die Zukunft.

<sup>36</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok

Ebenso kann über die Möglichkeit der Kategorisierung von Bildungs- und Ausbildungsinhalten erkannt werden, wer in der Lage ist, gewisse Ereignisse zu ermöglichen. Hat eine Einzelperson überhaupt die Chance gewisse Fähigkeiten zu erhalten oder kann nur eine Gruppe gewisse Dinge umsetzen?

In der Graphik Detail B wird die Sekundärinformation dargestellt, die um ein Ereignis herum auftritt. Aus diesen Sekundärinformationen lassen sich Ereignisse erklären, Planungen für Ereignisse ableiten, Motivationen erkennen, Wirkungen abschätzen und letztlich auch Lagebilder ableiten.



**Abbildung 22: Modell der zeitlichen Quelleninterdependenz  
Detail B: Sekundärinfo<sup>37</sup>**

Ein zentraler Faktor darin ist die Motivation von Akteuren zu einem Ereignis. Wer kann aus seinem Umfeld überhaupt Interesse an gewissen Dingen haben? Kulturelle oder politische, wirtschaftliche oder Neugier bedingte Motive lassen sich nach gewisser Zeit über Muster leichter wiederfinden. Diese Form der Informationsgewinnung und -strukturierung bildet dabei die Voraussetzung für eine effektive Analyse. Unterstützt durch die Erkenntnisse der Assoziationsheuristik ist es möglich gewesen

<sup>37</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok

Motivationshinweise zu bekommen und diese nachhaltig sichtbar zu machen! Die Rolle der Crowd Assoziationen bekommt hier besonderes Augenmerk und auch hier sind wissenschaftliche Fragestellungen äußerst relevant und in Zukunft muss eine systematische Erkennung von weak signals angedacht werden.

Im dritten Detail Diagramm zum Modell der zeitlichen Quellen Interdependenz wird deutlich, wie diese Information aus einer umfassenden Analyse von Ereignissen zum Horizon Scanning verwendet werden kann.

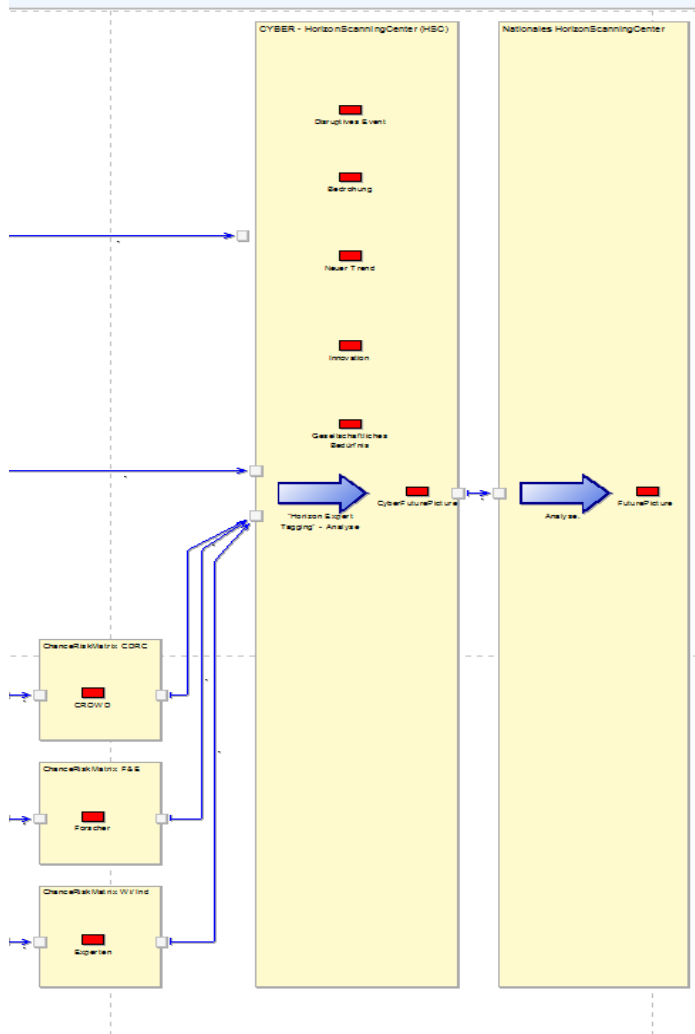


Abbildung 23: Modell der zeitlichen Quelleninterdependenz  
Detail C: Horizon Scanning<sup>38</sup>

Aus dem Ereignis Hintergrund, aus den Planungs- und Schulungsdokumenten, aus den Motivationen, den Interessen, den Fähigkeitsentwicklungen, den Zukunftserwartungen, den Strategiepapieren

<sup>38</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok



und den Visionen, lassen sich neue und noch unbekannte Ereignisse in ihrer Anbahnung erkennen.

Ausgewertet durch verschiedenen Experten oder Expertengruppen können die User des CDRC so auf ein aktuelles und umfassendes Informationspotential zurückgreifen, das täglich aus qualitätsgesicherten Quellen angereichert wird. Das an der ZentDok mit AIT als Kooperationspartner entwickelte Phasenmodell veranschaulicht auch in Details, wie zeitliche und inhaltliche Aspekte zu Grundlagen für eine Ereignisgeschichte werden. Ursache und Wirkung sollen und müssen dabei erkennbar bleiben. Sonst können unliebsame Wirkungen nie in ihrer Ursache erfasst und verhindert werden.

## **5.2 CYBER Informationsplattformen**

Um alle Anforderungen der Kategorisierung, Strukturierung, Metadatenerfassung, Typisierung und Aktualität für den Nutzer rasch sichtbar und abrufbar zu machen, wurden unterschiedliche Informationsplattformen evaluiert und einige Module zu einer eigenen Informationsplattform entwickelt. U.A. wurde eine sog. „Cyber Ushahidi Informationsplattform“ aufgesetzt und in Betrieb genommen. Damit kann nach allen Kriterien rasch ein Einblick in das „Tagesgeschehen“ in der Domain Cyber genommen werden. Ein Vergleich mit automatisierten Verfahren war nicht annähernd in der Lage die Arbeit der Crowd in ähnlicher Qualität zu liefern. Weder die Kategorisierung noch die Typisierung, und schon gar nicht das Erkennen der wichtigsten Zusatzinformationen auf einer Site („Site Info“) oder von neuen Zusammenhängen war für sog. „Crawler“ möglich.

Die folgende Graphik zeigt die Einstiegsseite der CDRC Cyber Informationsplattform, die auf Basis von Ushahidi umgesetzt wurde.



Abbildung 24: Startseite der CDRC Cyber Informationsplattform<sup>39</sup>

Alle administrativen und betrieblichen Erfordernisse der weltweit eingesetzten Open Source Plattform „Ushahidi“ konnten durch Kaderpersonal und Rekruten der ZentDok selbst erfüllt werden. Auch Neuanpassungen des Systems stellten im ersten Jahr des Betriebes kein Problem dar.

Ushahidi stellt dabei ein System zur Verfügung, welches für die Bedürfnisse des Crowdsourcings im Krisenfall entwickelt wurde. Durch die Anpassungen am CDRC ist es auch geeignet in der Cyber Domain als Meldesystem zu dienen, welches zumindest einfache Formen der Visualisierung der Ergebnisse bereit stellt.

<sup>39</sup> Quelle Screenshot: CDRC Informationsplattform

Der zentrale Vorteil bei der Auswahl von Ushahidi war, dass in dieser Umgebung Meldungen kategorisiert in das Informationssystem eingebracht werden können, dass dies einem Dokumententyp (DocType) der ursprünglichen Quelle zugeordnet werden können und dass Site Infos und andere Metadaten jeder Meldung mitgegeben werden können. Weiters ist es möglich gewesen das Expert Tagging in diesem System zu verwirklichen. Alle diese Elemente sind notwendig, um die Informationen aus der Cyberdomain, entsprechend dem Monitoring Konzeptes des CDRC, abgeleitet aus dem Modell der zeitlichen Quelleninterdependenz zur qualitätsgesicherten Informationsverdichtung im OSInfo Bereich (MQqI), zur Verfügung zu stellen.

Durch die Kategorisierung nach Thema und DocType wird es möglich, auf der Ushahidi Plattform erste zeitliche Zusammenhänge zwischen den unterschiedlichen DocTypes zu veranschaulichen.

Im Folgenden werden das CDRC Kategoriensystem, die CDRC DocTypes und die unterschiedlichen Formen von Metadaten im Informationsmanagement Prozess des CDRC vorgestellt.

### **5.2.1 CYBER – Kategoriensystem**

Das Cyber Kategoriensystem des CDRC dient dazu schon bei der Meldung in das System die Meldung grob inhaltlich zu clustern. Dadurch wird es für den User möglich, Meldung nach Kategorien zu filtern und beim Bedarfsfalle Meldung nur zu einzelnen Kategorien zu beobachten und zu bewerten. Sollte z. B. ein Cyber Experte sich nur für Angriffsmethoden aus der technischen Perspektive interessieren, so kann er alle anderen Meldungen ausblenden und bekommt nur diese über das System zur Gesicht.

Wie aus der Graphik ersichtlich, werden die Meldungen derzeit nach 29 Kategorien klassifiziert.

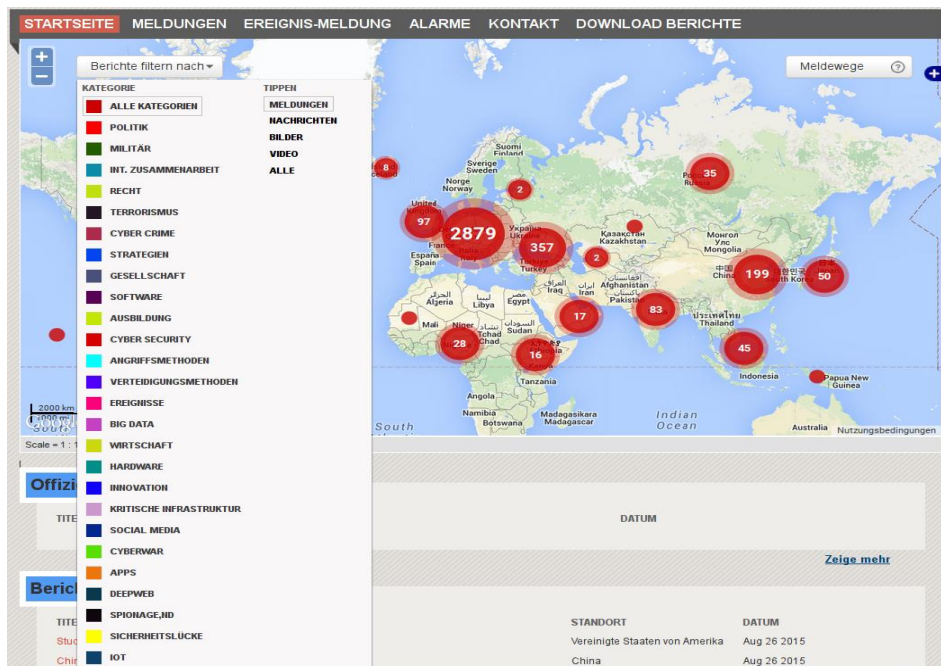


Abbildung 25: Kategoriensystem der CDRC Cyber Informationsplattform<sup>40</sup>

Diese Kategorien sind aus der praktischen täglichen Arbeit entstanden und unterliegen einem stetigen Wandlungs- und Verbesserungsprozess. Sollte sich bei Crowd Recherchen zeigen, insbesondere von HPS, dass neue Kategorien notwendig sind, so werden diese in das System aufgenommen. Sollte zu einzelnen Kategorien zu wenig Meldungen auftauchen oder diese nicht zuordenbar sein, so wird die Kategorie wieder aus dem System herausgenommen.

Im Laufe der Zeit ist es zu einer sukzessiven Stabilisierung des Kategorisierungssystems gekommen, so dass diese zunehmend seltener verändert werden müssen. Mit der Zeit muss sich erweisen, dass obenstehende Kategorien auch für die User passen. Aus der Nutzerstatistik wird im Laufe der Zeit ersichtlich, welche Kategorien mehr nachgefragt werden und welche weniger. Durch die Adaption wird das

<sup>40</sup> Quelle Screenshot: CDRC Informationsplattform

Informationssystem für die User zunehmend leichter zu nutzen und damit wertvoller.

### **5.2.2 CYBER - DocType, Site Info, Meta Data, Expert Tagging**

Von den vielen Metadaten, die im Crowd OSInfo Prozess des CDRC entstehen, spielt vor allem der Dokumententyp eine wesentliche Rolle. Wie schon bei der Vorstellung des Modells der zeitlichen Quelleninterdependenz zur qualitätsgesicherten Informationsverdichtung im OSInfo Bereich (MQqI) erwähnt, können die Dokumententypen (DocTypes) dazu verwendet werden, um einen zeitlichen Zusammenhang zwischen Informationen aus unterschiedlichen Quellen herzustellen. Ushahidi unterstützt bis zu einem gewissen Grade diese Form von Auswertungen, indem nach DocTypes gefiltert werden kann und diese gefilterten Informationen dann in weiterer Folge in einer Zeitleiste dargestellt werden können. Zeitleisten mit unterschiedlichen DocTypes können verglichen werden. Die folgende Graphik zeigt die DocTypes vom CDRC Ushahidi und eine entsprechende zeitliche Sortierung.

Zeige Berichte von Jan 05, 2008 bis Aug 26, 2015 [Datumsbereich ändern](#)

Liste Karte 1 2 3 4 5 6 ... 143 144 1-50 of 7166 Meldungen < >

**Study shows that internet search engines have the power to swing elections.** 10:48 Aug 26, 2015

As a society, we are happily ensconced in the internet era. And we're sure that you, oh wonderful blog readers, are among the first to...

**China Caught Spying on Tibetan Activists and Neighbouring Countries** 10:42 Aug 26, 2015

China has been a usual suspect behind the cyber attacks carried out worldwide — From U.S. to India everyone blames the red army for...

**Prisoner Hacked Jail TV System To Watch and Share Porn** 10:26 Aug 26, 2015

A high-security Norwegian prison has experienced a unique event: a man was caught watching pornography on his TV after he hacked into the...

**Office 2016 kommt offenbar am 22. September** 08:52 Aug 26, 2015

Eine Quelle von WinFuture will diesen Termin von einer Intranet-Seite erfahren haben, über die Microsoft seine Angestellten auf den...

**Jugend-Vertrauen: Snapchat glänzt vor LinkedIn** 08:46 Aug 26, 2015

User finden befreundeten Vorgesetzten zunehmend unproblematischer Online. Unterschiedliches Vertrauen verwirrt (Foto: Andreas...

**Verhetzung auf Facebook: Acht Monate unbeding für Welsler** 08:45 Aug 26, 2015

Ein 28-Jähriger hat am Dienstag am Landesgericht Wels acht Monate unbeding wegen Verhetzung ausgefasst, weil er auf Facebook zur Vergasung...

**"Autonome Maschinen werden unweigerlich Schaden anrichten"** 08:22 Aug 26, 2015

Damit die Roboter der Zukunft enger mit Menschen kooperieren können, müssen sie viel über uns wissen. Matthias Scheutz sucht...

**Alibi-Router sollen Daten um Zensurzonen herumleiten** 08:18 Aug 26, 2015

Nutzer haben gegenwärtig wenig Kontrolle darüber, über welche Regionen in der Welt ihre Daten geleitet werden. Das will...

**Günstige bionische Hand gewinnt Designpreis** 08:17 Aug 26, 2015

**Ashley Madison wusste von gravierenden Sicherheitsmängeln** 08:15 Aug 26, 2015

Einige Wochen vor dem Angriff des Impact Teams warnten interne Sicherheitsexperten vor gravierenden Mängeln in der Infrastruktur...

**Regierung soll Telekom exklusives Vectoring erlaubt haben** 08:13 Aug 26, 2015

Der Telekom wird offenbar ihr exklusives Vectoring am Hauptverteiler genehmigt

**Berichte filtern nach**

Kategorie [Löschen](#)

Standort [Löschen](#)

Tippen [Löschen](#)

Medien [Löschen](#)

Bestätigung [Löschen](#)

**Eigene Felder** [Löschen](#)

**Sprache \***

---Nicht ausgewählt---

**Erscheinungsdatum**

**Autor**

**DocType \***

News-Innovativ  PATENTE

Forschungs und Entwicklungs Programme  Forschungs und Entwicklungs Projekte

Forschungs und Entwicklungs Plattformen  STATE of the ART Plattformen

Konferenzen  Dissertationen-GL

Ausbildung  Produkte

Ereignisse  Zukunft

WEAK-Signals  strategische Infos

Trendbrüche  CYBER-Plattform

**Site-Info**

Link  Org

Al:teur  Blog/SocMedia

Produkt  Werbung

Event

**Hoizon Expert Tagging - CYBER**

Disruptives Event  neuer Trend

Bedrohung  Innovation

gesellschaftliches Bedürfnis

[Alle Filter zurücksetzen](#) **BERICHTE FILTERN**

Abbildung 26: Dokumententypen der CDRC Cyber Informationsplattform<sup>41</sup>

Neben den DocTypes bietet auch das Expert Tagging einen Ansatz der Informationsverdichtung und Anreicherung. Das Expert Tagging System bietet eine Möglichkeit, jeder Meldung im System eine inhaltliche Markierung eines Experten zuzuordnen. So sind z.B. Klassifikationen, wie neuer Trend, Innovation, disruptives Event, Bedrohung oder

<sup>41</sup> Quelle Screenshot: CDRC Informationsplattform

gesellschaftliche Bedürfnisse wertvoll für die Analyse von Horizon Scanning Experten. Diese können aus neuen disruptiven Events z.B. auf Veränderungen bei Trends schließen. Die Expert Tags können für die jeweilige Domain sehr rasch angepasst werden und sind ein wesentlicher Bestandteil für eines „Horizon Scanning Center“.

Neben den Tags werden noch eine ganze Reihe weiterer Metainformationen bei den Meldungen im Ushahidi System erhoben. Diese dienen dazu, dass später in nachfolgenden Analysen alle notwendigen Informationen zur Erstellung von Lagebildern vorhanden sind.

### **5.2.3 News Erfassungssystem LAMARR der CDRC Cyber Informationsplattform**

Für die Erfassung von Meldungen aus dem Internet, vor allem bei HPS, wurde ein eigenes System (LAMARR) entwickelt wo gewährleistet ist, dass ausschließlich relevanter Content ohne Werbung und sonstiger unnötiger Informationen in das Cyber Repository gelangt. Dabei werden standardisierte Metainformationen teilautomatisiert ausgelesen, zusätzliche Kategorisierungen durchgeführt und weiters auch automatisiert ein einheitlich formatiertes PDF erzeugt und im Repository gemeinsam mit den Metadaten und den Kategorien gespeichert. Damit soll ein qualitativ hochwertiges IR ermöglicht werden bzw. die Basis für professionelle Analyseprozesse geschaffen werden.

Startseite Meldungen Logout

**CDRC**  
Cyber Documentation & Research Center  
The owl is watching you...

Neue Meldung

Meldung bestätigen

Bestätigen der Meldung

Die Daten für den Artikel konnten erfolgreich geholt werden. Bitte überprüfen Sie die Daten und ergänzen Sie gegebenenfalls fehlende Daten und senden sie das Dokument dann ab.

**Titel der Meldung\***  
Schlagabtausch der Großmächte im "Cyber-War"

**Subtitel der Meldung**

**Kategorien laut Seite**

**Beschreibung\***  
Trojanersuite Duqu 2.0 gegen UN-Verhandlungen mit dem Iran, US Verwaltung öffentlicher Dienst verliert vier Mio Personalakten, deutscher Bundestag muss 20.000 PCs tauschen.  
Der Außenminister der Schweiz eröffnete Duqu 2.0 ist

**Erscheinungsdatum**  
09.06.2015

**Autor**  
Erich Möchel

**Ort\***  
USA

**Kategorien\***  
Cyber Crime Cyber Security Kritische Infrastruktur

**Doc-Type\***  
News-Innovativ  
Ausbildung Cyber-Plattform Dissertationen-GL Ereignisse  
Forschung und Entwicklung - Ausschreibungen  
Forschung und Entwicklung - Plattformen  
Forschung und Entwicklung - Programme  
Forschung und Entwicklung - Projekte Forschungsorganisation  
Konferenzen Patent Produkt Strategische Infos  
State of the Art Trendbrüche WEAK-Signals Zukunft

**Site-Info**  
Akteur Blog/Social Media Event Link Organisation  
Produkt Werbung

Abbildung 27: News Erfassungssystem LAMARR<sup>42</sup> der CDRC Cyber Informationsplattform

<sup>42</sup> Quelle Screenshot: CDRC Informationsplattform



## 5.2.4 Recherche Datenbank

Neben dem Customizing von Open Source Systemen wurden in der Pilotphase des CDRC auch eigene Systeme entwickelt. Die Recherche Datenbank mit dem dazugehörigen Meldesystem dient im CDRC dazu, Meldungen so einfach und effizient wie möglich zu gestalten.

Entworfen, entwickelt und betrieben durch Cyber Rekruten, ermöglicht die CDRC-Recherchedatenbank eine systematische Dokumentation, Bearbeitung und Analysemöglichkeit aller Inhalte und Metadaten der Recherchen des CDRC. In weiterer Folge wird auch eine Update-Systematik implementiert, die in regelmäßigen Abständen über „Stored Queries“ kategorisierte Recherchen wiederholt und aktualisiert.

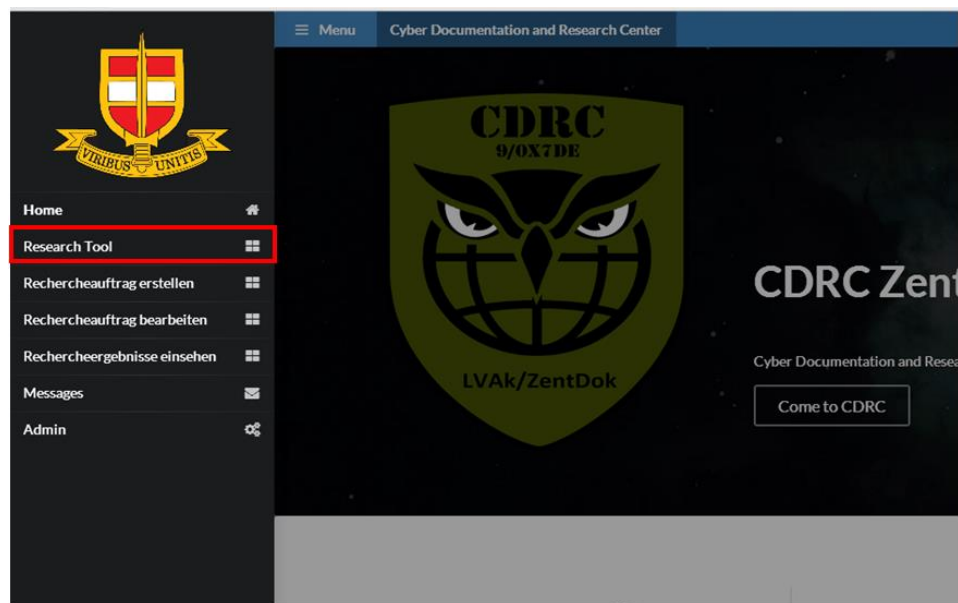


Abbildung 28: Recherche Datenbank der CDRC Cyber Informationsplattform<sup>43</sup>

Dazu wurden spezifische Deep Web Datenretrieval Agents entwickelt, die den Content von HPS extrahieren und sauber, mit Metadaten angereichert, in das CDRC Informationsmanagement bringen. Diese Retrieval Agents können in weiterer Folge auch dazu verwendet werden zuzunehmen

<sup>43</sup> Quelle Screenshot: CDRC Recherche DB

automatisiert die Daten aus dem Internet zu beschaffen. Es ist eine andauernde Forschungsfrage des CDRC festzustellen, welche Prozesse automatisiert werden können und welche besser mit menschlichen Fähigkeiten zu erledigen sind.

### 5.2.5 Multi Media Demonstrator

Durch die Beteiligung an KIRAS-Sicherheitsforschungsprojekten wurden uns durch den Partner SailLabs ein Multi Media Demonstrator für Test- und Entwicklungszwecke zur Verfügung gestellt. Damit konnten additiv zu den Inhalten des CDRC multimediale Inhalte dokumentiert und analysiert werden. Dies geschieht bei dem Demonstrator meist automatisiert und abgestützt auf modernsten Technologien.



Abbildung 29: Multimedia Demonstrator der CDRC Cyber Informationsplattform<sup>44</sup>

<sup>44</sup> Quelle Screenshot: Multimedia Demonstrator

Damit war es möglich, Quantität und Qualität mehrerer Verfahren der Medienanalyse und deren Ergebnisse zu untersuchen sowie zu vergleichen. Dies wird in weiterer Folge besondere Bedeutung erlangen, um zu versuchen, bei großen Mengen an Informationen eine hohe Qualität der Verarbeitung zu gewährleisten. In Zukunft wird es immer schwieriger werden in kurzer Zeit die besten Ergebnisse zu erzeugen, die einen hohen Grad an Vollständigkeit der jeweiligen Informationspotentiale erkennen lassen.

Neben der Anwendung bestehender Tools zur automatisierten Verarbeitung multimedialer und multilingualer Informationen wird deswegen stetig an neuen Lösungen zur Verbesserung der automatisierten Informationsverarbeitung gearbeitet und in Kooperationen mit externen Partnern, wie dem AIT, auch geforscht.

### **5.2.6 Terminologiesystem – ProTerm**

Eine Basis für die Automatisierung der Informationsverarbeitung im CDRC stellen Tools zur Terminologiebearbeitung dar. Das seit Jahren in der ZentDok mit Partnern entwickelte und eingesetzte Terminologie-Werkzeug ProTerm wurde im CDRC verstärkt eingesetzt und weiterentwickelt.

ProTerm ist ein Werkzeug für Terminologiearbeit & Textanalyse. Es eignet sich weiters für die Organisation & Aktualisierung verwendeter Terminologie, die Extraktion von Benennungen aus verwendeten Dokumenten, zur Translationsunterstützung, für die Integration in bestehender Thesauri und damit für eine Anpassung bereits vorhandener Terminologien. Auch die Beschlagwortung von Informationseinheiten und eine kontrollierte Suchmethode werden durch ProTerm unterstützt.

Die Integration neuer Sprachen zeigt sich in Zukunft als besondere Herausforderung. Unten angeführte Beispiele veranschaulichen erste Anwendungen in Türkisch, wobei weitere Sprachen je nach Verfügbarkeit der Sprachenkompetenz wie Polnisch und Russisch bearbeitet wurden. In all diesen Sprachen wurden nicht nur neue Begriffe gefunden, sondern auch über Vergleichstexte Unterstützungsleistungen für Translationsaufgaben erbracht.

Die folgende Graphik zeigt eine Liste aktueller Cyber Terminologien in der Verarbeitung mit ProTerm.

Benennung	Status	Doc	Max	Zlen	Len	Normiert
siber	acc	33	1795	5	1	2
Güvenlik	acc	32	1038	9	1	2
bilgi	acc	24	700	8	1	0
bag	stop	32	682	4	1	0
bilgi	acc	30	519	5	1	2
sup	acc	16	414	4	1	2
bilgiyas	acc	20	403	10	1	2
Ceza	acc	9	302	4	1	0
Bag	acc	28	297	4	1	0
ulusal	acc	17	264	6	1	0
internet	acc	22	259	8	1	2
saldn	acc	24	259	9	1	2
TUBITAK	acc	12	251	9	1	0
Security	acc	17	240	8	1	0
veri	acc	18	223	4	1	2
ULUSLARARASI	acc	18	222	12	1	0
karpi	acc	26	218	7	1	0
supun	acc	5	214	6	1	0
suçlar	acc	13	210	7	1	0
elektronik	acc	15	208	10	1	0
Suçlan	acc	7	189	9	1	0
internet	acc	14	178	9	1	0
Sevrima	acc	22	178	7	1	0
trajika	acc	4	172	6	1	0
Univeritesi	acc	9	170	13	1	0
Cyber	acc	14	169	5	1	0
Kamu	acc	18	163	4	1	0
zans	acc	17	157	5	1	0
supu	acc	7	157	5	1	0
IP	acc	12	153	2	1	0
madde	acc	8	145	5	1	2
ka	stop	20	139	2	1	0

Abbildung 30: Neue Terminologie in ProTerm (türkisch)<sup>45</sup>

Neben Worthäufigkeiten, Wortzusammenhängen und Synonymen lassen sich damit auch Normalisierungen und Übersetzungslisten pflegen, sodass diese bei automatisierten Systemen des Text Minings eingesetzt werden können.

Die nachfolgende Graphik zeigt einen Ausschnitt der Übersetzungsterminologie in Englisch und Türkisch.

<sup>45</sup> Quelle Screenshot: ProTerm – CDRC Datensatz

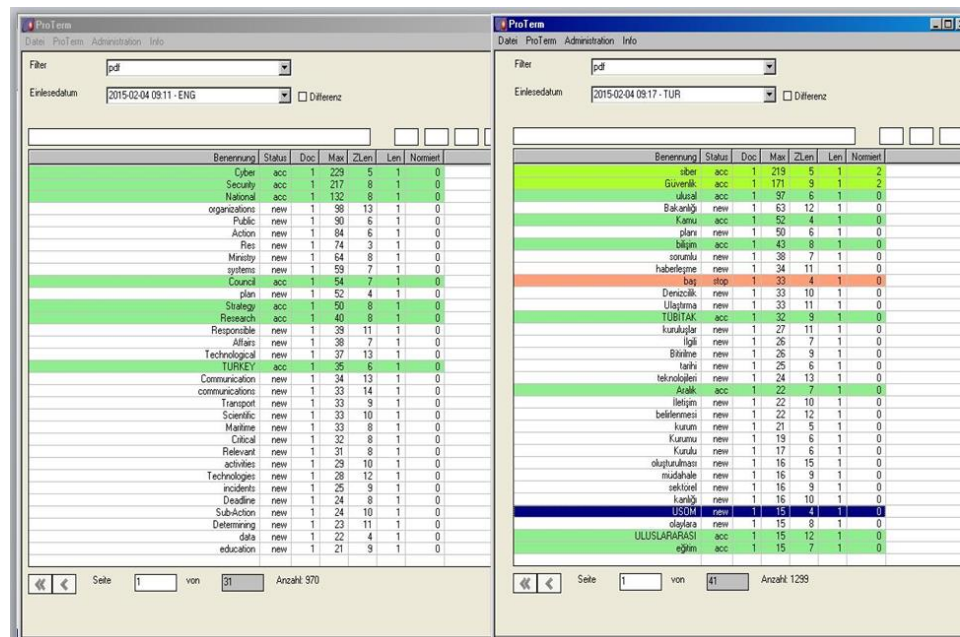


Abbildung 31: TransTerm Übersetzungsterminologie (Englisch - Türkisch)<sup>46</sup>

Aus den Wortlisten in Englisch und Türkisch wird ersichtlich, dass diese als Grundlage für einen Multilingualen Cyber Thesaurus dienen können. Bei den Crowd Recherchen des CDRC hat sich gezeigt, dass es weder aus der Perspektive staatlicher Services, noch aus der Perspektive von Forschung und Industrie ausreichend ist, sich auf den deutschen und englischsprachigen Content alleine zu beschränken. Vielmehr ist es so, dass z.B. in türkischer Sprache wertvolle Informationen über z.B. über die Kooperation China-Japan im Cyber Security Bereich veröffentlicht wurden.

Deswegen ist es für ein vollständiges und aktuelles Lagebild im Cyber Bereich notwendig, das Internet in den unterschiedlichsten Sprachen zu betrachten. Idealerweise lassen sich diese daraus gewonnenen Informationen so verwerten, dass die Informationen eine Entscheidungsunterstützung bieten in Bezug eine kurz, mittel und langfristige Planung. Eine Institutionalisierung all der Prozesse im CDRC soll langfristig dazu

<sup>46</sup> Quelle Screenshot: ProTerm – CDRC Datensatz

führen, dass das CDRC die Aufgaben eines Horizon Scanning Centers erfüllen kann.

## 6 Horizon Scanning Center „CYBER“

In den letzten Jahren werden weltweit vermehrt Horizon Scanning Center aufgebaut, um Informationen zu liefern, die staatliche Einrichtungen bei der langfristigen Planung unterstützt. Ziel dieser Horizon Scanning Center ist es, möglichst frühzeitig und möglichst vollständig alle planungsrelevanten Informationen für eine bestimmte Domain zu sammeln und durch Experten bewerten zu lassen.

Eine Beschreibung von Zentren in Singapur, den Niederlanden, bei den UN und in der EU findet sich in der Publikation „WM im ÖBH - Foresight in der strategischen Langfristplanung“, in der Schriftenreihe der LVAk<sup>47</sup>.

In der nachfolgenden Abbildung wird ein Prozessdiagramm vorgestellt, welches die Bearbeitungsschritte veranschaulicht, die notwendig sind, um in der Domain Cyber eine Möglichkeit für „Horizon Scanning“ aufzubauen. Durch diese erste Festlegung wurden alle Vorteile der Arbeitsweise mit der Crowd in das Gesamtsystem eingebettet. Die Kombination mit qualitätsgesicherten automatisierten Informationsgewinnungsverfahren stellt sich als weitere zukünftige Herausforderung dar. Dies gilt im besonderen Maße auch für hybride und mehrsprachige Informationselemente wie bspw. Videos und Social Media. Alle Möglichkeiten für eine kontinuierliche Weiterentwicklung, auch für andere Domains, können so systematisch erfasst, getestet und erweitert werden.

---

<sup>47</sup> Vgl. Göllner, J., Klerx, J., Mak, K., Meurers, C.; „Wissensmanagement im ÖBH - Foresight in der strategischen Langfristplanung“, Band 5/2015, Wien, 2015, herausgegeben von Göllner, J., Klerx, J., Mak, K.

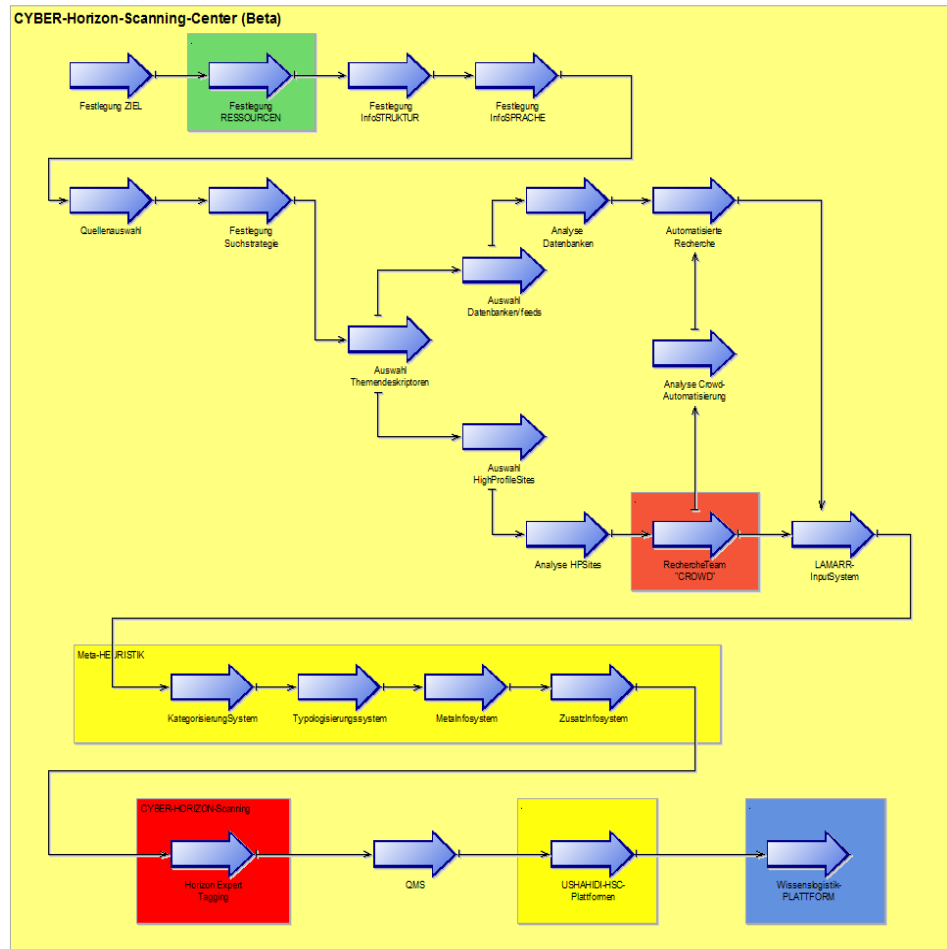


Abbildung 32: Prozessdiagramm eines Horizon Scanning Centers<sup>48</sup>

Nach der Festlegung von Ressourcen, notwendiger Infrastruktur und relevanter Sprachen ist für ein Horizon Scanning vor allem auch die Quellenauswahl und die Suchstrategien ein zentrales Element. Beides sollte in einem iterativen Prozess erfolgen, so dass sowohl die Quellenauswahl, als auch die Suchstrategie sich über die Zeit verbessern.

<sup>48</sup> Quelle: eigene Darstellung, mit WM-Werkzeug PROMOTE® der ZentDok

Neben den Quellen für die Crowd OSInfo Analyse werden auch Deskriptoren, Datenbanken und Feeds für die Expertenrecherche benötigt. Aus dem Zusammenspiel dieser beiden Recherchemethoden ergibt sich bei der Aufarbeitung ein erstes Informationslagebild.

Einerseits stellt sich zu diesem Zeitpunkt schon die Frage nach dem Grade der Automatisierung. Auf der anderen Seite müssen aus den vorgefundenen Informationen haltbare Analysen erzeugt werden, die je nach Ziel des Zentrums auf zukünftige Szenarien hinweisen.

Szenarien in diesem Sinne stellen nachvollziehbare Zukunftserwartungen dar, die transparent begründet und bewertet werden, so dass bei einer Veränderung des Kenntnisstands deutlich wird, wie dieser zu veränderten Szenarien führt. Alle Zukunftsbewertungen, wie z.B. erwünschte und unerwünschte Szenarien, mögliche Bedrohungen, mögliche disruptiven Events, mögliche Extremevents, und mögliche Trends werden, soweit möglich und sinnvoll, mit Eintrittswahrscheinlichkeiten und Impact Kenngrößen bewertet, um sie einer Risikoanalyse zugänglich zu machen.

Durch die kontinuierliche Vorbereitung und Aufbereitung ist es möglich, im Einzelfalle, auch bei unerwarteten Ereignissen, schnell zu reagieren. Diese Fähigkeit der erweiterten Resilienz wird allerdings damit erkauft, dass die Ressourcen für das kontinuierliche Horizon Scanning bezahlt werden müssen. D.h. der gesamtstaatliche Nutzen liegt darin, dass die Unsicherheit bei möglichen zukünftigen Entwicklungen verringert wird, zugunsten einer besseren Planbarkeit und einer besseren Vorbereitung auf unerwünschte Entwicklungen.

Im speziellen Falle des CDRC hat sich schon jetzt ein zusätzlicher gesamtstaatlicher Nutzen in mehrfacherweise gezeigt. Die Innovationsgeschwindigkeit im Cyber Bereich ist derart hoch, dass klassische Ausbildungsmechanismen und klassische Formen der Forschung und Lehre versagen. Mit den Methoden des CDRC wird sowohl allen Mitgliedern am Crowdsourcing, als auch allen Experten im Umfeld des CDRC eine völlig neue Sicht auf die Entwicklungen im Cyber Bereich gegeben. Die Methoden einen State of the Art zu einem bestimmten Zeitpunkt zu definieren haben sich durch die Innovationen des CDRC nachhaltig verändert. Expertenrecherchen ohne Unterstützung durch eine



Crowd Recherche sind einfach qualitativ schlechter, so dass vermutlich zunehmend ergänzende Crowd Recherchen mit eingesetzt werden.

Mit den methodischen Innovationen im CDRC wird es in Zukunft leichter, mit der Innovationsgeschwindigkeit im Cyber Bereich umzugehen. Diese Erfahrung nehmen wohl alle am Pilotprojekt des CDRC Beteiligten mit.

## **7 Ausblick**

Die Pilotphasen des CDRC haben gezeigt, dass Methoden wie die Crowd Recherche und das Modell der zeitlichen Quelleninterdependenz zur Informationsverdichtung mächtige Werkzeuge sind, um Informationsräume, wie z.B. den Cyber Raum strategisch aufzubereiten. Diese Methoden könnten sowohl für staatliche Services, als auch für die Forschung und die industrielle Entwicklung die Informationslage deutlich verbessern.

Wissensintensive Produkte, wie eine Forschungsdatenbank „Cyber“, eine Wissenslogistik „Cyber“ basierend auf den Wissenslogistik Modellen des CDRC können dazu beitragen, dass ein sich schnell entwickelndes Gebiet, wie die Cyber Domain, wieder einer koordinierten Langfriststrategie zugänglich wird. Nachdem die Digitalisierung in Deutschland als wichtigstes strategisches Zukunftsfeld identifiziert wurde, sind Methoden zum Umgang mit hoher Innovationsgeschwindigkeit und einer Wissenschaft nach der digitalen Revolution wichtiger denn je.

Ein kontinuierliches Horizon-Scanning führt nicht nur dazu, dass man von Neuigkeiten am Horizont nicht überrascht wird. Es trägt auch dazu bei, dass die Fähigkeiten des Horizon Scanning und des digitalen Wissensmanagements stetig weiter entwickelt werden. Diese Fähigkeiten tragen dazu bei, dass „Cyber Lagebilder“ immer in einer aktuellen Version verfügbar sind. Sie tragen auch dazu bei, dass Foresight Informationen über mögliche langfristige zukünftige Entwicklungen zur Verfügung stehen, wenn diese z.B. in der Forschungsplanung benötigt werden. Die nachfolgende Graphik fasst die wesentlichen Elemente des Modells der zeitlichen Quelleninterdependenz zur Informationsverdichtung aus Kapitel 5.1 zusammen. Die kontinuierliche Beobachtung des Ereignisraumes der

Cyber Domain lässt nur einen sehr begrenzten Rückschluss auf Entwicklungstendenzen im Cyber Raum zu. Der Einbezug von historischen Informationen und Foresight Informationen erweitert diese Perspektive ungemein.

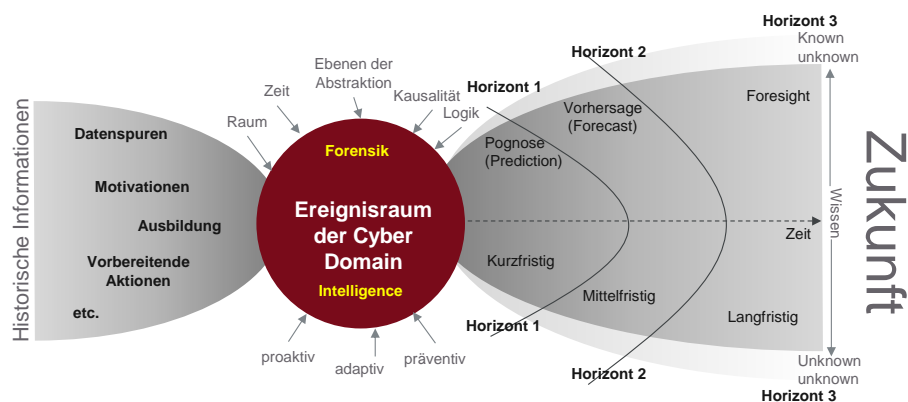


Abbildung 33: Meta-Logik zur Entwicklung proaktiver and adaptiver Strategien<sup>49</sup>

Ohne adäquate technische Unterstützung ist diese Erweiterung der Perspektive nicht möglich. Methoden der Textanalyse „Cyber“, der semiautomatischen Translation, der mehrsprachige Thesauri „Cyber“ und moderner High Performance Crawlingsysteme ermöglichen erst ein effizientes Arbeiten der Crowd. Dabei ist die Entwicklung keineswegs abgeschlossen. Viele der aktuellen Prozesse benötigen wissenschaftliche Instrumente die noch entwickelt werden müssen.

<sup>49</sup> Quelle eigene Darstellung

Deswegen bieten sich F & E Aktivitäten zur

1. automatischen Textanalyse
2. semiautomatischen Translation mit Erstellung mehrsprachiger Thesauri
3. Entwicklung von Big Data fähigen Demonstratoren für High Performance „Cyber“ Crawlingsystemen
4. neuen und optimierten „Crowd-Recherche-Modell“
5. sowie zu „Crowd-Auswerte-Modell“

an.

Bis ein Horizon Scanning Center betrieben werden kann, welches alle in der Abbildung 32 zusammengefassten Funktionen erfüllt, wird wohl noch einige Zeit vergehen. Die derzeitigen Ergebnisse des CDRC deuten allerdings jetzt schon an, dass sich mit diesen Prozessen die Open Source Analysen und das wissenschaftliche Wissensmanagement ändern werden.

## 8 Literatur

- Expert Systems with Applications, Volume 42, Issue 11, 2015,
- Journal of Surgical Research, Volume 187, Issue 1, 2014
- Alemu, G., Stevens, B., The Web 2.0 paradigm and the emergence of socially-constructed metadata approaches, An Emergent Theory of Digital Library Metadata, 2015
- Göllner, J., Klerx, J., Mak, K. (Hrsg.); „Wissensmanagement im ÖBH - Foresight in der strategischen Langfristplanung“, Schriftenreihe der Landesverteidigungsakademie, Band 5/2015
- Göllner, J., Meurers, C., Quirchmayr, G. (Hrsg.), „Wissensmanagement im ÖBH – Einführung in die Informationslogistik als Grundlage zur Wissens- und Organisationsentwicklung“, Schriftenreihe der Landesverteidigungsakademie, Band 20/2015, Wien, 2015
- Applied Geography, Volume 35, Issues 1–2, November 2012
- Wilby, Blachfellner, Hofkirchner (Eds.), Book of Abstracts, EMCSR-European Meetings on Cybernetics and Systems Research, Wien, 2014;
- Mak, K., Woitsch, R., „Der Einsatz des prozessorientierten WM-Werkzeuges PROMOTE® in der ZentDok der LVak“, Schriftenreihe der Landesverteidigungsakademie, 19/2005
- New Scientist, Volume 218, Issue 2915, 2013
- International Journal of Intelligence and CounterIntelligence, Volume 28, Issue 3, 2015
- Procedia Computer Science, Volume 32, 2014
- Healthcare, Volume 3, Issue 1, 2015
- Government Information Quarterly, Volume 29, Issue 1, 2012
- Biological Conservation, Volume 181, January 2015

**Web Quellen:**

Mashable, URL:

<http://mashable.com/2011/04/01/make-money-crowdworking/>,  
[8.9.2015]

International News Analysis Today, URL:

<http://inatoday.com/terror%20goes%20analytical%20032604.htm>,  
[8.9.2015]

The Daily Beast, URL:

<http://www.thedailybeast.com/articles/2014/10/01/isis-is-winning-the-online-jihad-against-the-west.html>, [23.09.2015]

## 9 Abbildungsverzeichnis

Abbildung 1: Positionierung des CDRC im Organigramm der ZentDok/LVAk.....	8
Abbildung 2: Integration des CDRC in die Prozessstruktur der ZentDok	11
Abbildung 3: Wissensentwicklungsmechanismus: Z-Modell.....	13
Abbildung 4: Produktlandkarte des CDRC in der ZentDok/LVAk.....	15
Abbildung 5: Prozessdiagramm Cyber-Relevanz Monitoring.....	16
Abbildung 6: Prozessdiagramm Cyber-Dokumentation.....	18
Abbildung 7: Prozessdiagramm Cyber-Fachinformation .....	20
Abbildung 8: Prozessdiagramm Cyber-Research .....	22
Abbildung 9: Übersicht Crowd Assoziationsheuristik.....	27
Abbildung 10: Übersicht Crowd Suche Fragenheuristik .....	31
Abbildung 11: Prozessdiagramm Crowd Recherche.....	34
Abbildung 12: Prozessdiagramm Crowd Recherche Teil 1: Crowd.....	36
Abbildung 13: Prozessdiagramm Crowd Recherche Teil 2: Zusammenführung.....	38
Abbildung 14: Kriterienkatalog von High Profile Sites .....	40
Abbildung 15: CROWD – Assoziations-Matrix-Prozess („Chance – Risiko“) .....	43
Abbildung 16: Chancen Risiko Matrix im Bereich Cyber .....	45
Abbildung 17: Prozessdiagramm Integration Experteneinschätzung/ Crowd Recherche.....	47
Abbildung 18: Prozessdiagramm zur Zusammenführung Detail A .....	48
Abbildung 19: Prozessdiagramm zur Zusammenführung Detail B.....	49
Abbildung 20: Modell der zeitlichen Quelleninterdependenz zur Informationsverdichtung.....	51
Abbildung 21: Modell der zeitlichen Quelleninterdependenz Detail A: Infogewinnung.....	52
Abbildung 22: Modell der zeitlichen Quelleninterdependenz Detail B: Sekundärinfo.....	53
Abbildung 23: Modell der zeitlichen Quelleninterdependenz Detail C: Horizon Scanning.....	55
Abbildung 24: Startseite der CDRC Cyber Informationsplattform.....	57
Abbildung 25: Kategoriensystem der CDRC Cyber Informationsplattform.....	59

Abbildung 26: Dokumententypen der CDRC Cyber Informationsplattform .....	61
Abbildung 27: News Erfassungssystem LAMARR der CDRC Cyber Informationsplattform .....	63
Abbildung 28: Recherche Datenbank der CDRC Cyber Informationsplattform .....	64
Abbildung 29: Multimedia Demonstrator der CDRC Cyber Informationsplattform .....	65
Abbildung 30: Neue Terminologie in ProTerm (türkisch).....	67
Abbildung 31: TransTerm Übersetzungsterminologie (Englisch - Türkisch).....	68
Abbildung 32: Prozessdiagramm eines Horizon Scanning Centers .....	70
Abbildung 33: Meta-Logik zur Entwicklung proaktiver and adaptiver Strategien .....	73

## 10 Autoren

Ing. Mag. Klaus MAK  
Oberst des höheren militärfachlichen Dienstes  
Leiter der Zentraldokumentation an der Landesverteidigungsakademie  
klaus.mak@bmlv.gv.at

Dr. Joachim KLERX  
Forscher im Innovation System Department des Austrian Institute of  
Technology (AIT)

Hans Christian PILLES, ADir RgR  
Leiter Technische Dokumentation  
an der Zentraldokumentation/Landesverteidigungsakademie

Dipl.-Ing. Johannes GÖLLNER, M.Sc.  
Leiter Referat Wissensmanagement  
an der Zentraldokumentation/Landesverteidigungsakademie