



Användartest med datorverktyg för multihypoteshantering

MAGDALENA GRANÅSEN, PETER SVENMARCK, JOHANNA VON FEILITZEN,
ROBERT FORSGREN, TOVE GUSTAVI, MAJA KARASALO

Magdalena Granåsen, Peter Svenmarck, Johanna von Feilitzen, Robert Forsgren, Tove Gustavi, Maja Karasalo

Användartest med datorverktyg för multihypoteshantering

Titel	Användartest med datorverktyg för multihypoteshantering
Title	User test with computer tool for analysis of multiple hypotheses
Rapportnr/Report no	FOI-R--3918--SE
Månad/Month	Juni
Utgivningsår/Year	2014
Antal sidor/Pages	41 p
ISSN	1650-1942
Kund/Customer	FMV
Forskningsområde	1. Beslutsstödssystem och informationsfusion
FoT-område	Ledning och MSI
Projektnr/Project no	E34218
Godkänd av/Approved by	Lars Höstbeck
Ansvarig avdelning	Informations- och aerosystem

Detta verk är skyddat enligt lagen (1960:729) om upphovsrätt till litterära och konstnärliga verk. All form av kopiering, översättning eller bearbetning utan medgivande är förbjuden

This work is protected under the Act on Copyright in Literary and Artistic Works (SFS 1960:729). Any form of reproduction, translation or modification without permission is prohibited.

Sammanfattning

Hyppo är ett verktyg för multihypoteshantering som utvecklats av FOI. Rapporten beskriver metod och resultat från två användartester med syfte att studera verktygets användbarhet och nytta samt att identifiera kortsiktiga och långsiktiga utvecklingsbehov för kommande versioner.

Hyppo kombinerar de två analytiska metoderna *morfologisk analys* och *analys av konkurrerande hypoteser* (ACH). Morfologisk analys används ofta för att konkretisera och analysera möjliga händelseutvecklingar för så kallade *vildvuxna problem*. Under den morfologiska analysen skapas en morfologisk modell som består av en matris med ett antal dimensioner och möjliga tillstånd för varje dimension. I Hyppo används denna matris (i Hyppo kallas detta för det morfologiska fältet) till att generera hypoteser för det aktuella problemet. En hypotes i Hyppo består av ett tillstånd från varje dimension. Ett enkelt exempel kan vara hypoteser relaterade till en mordutredning, där *gärningsman*, *mordvapen* och *motiv* identifierats som viktiga dimensioner. En konkret hypotes kan då vara ”Gärningsmannen *Lee Harvey Oswald* med mordvapnet *Colt Holster Pistol* och motivet *politiskt motiverad terrorism*”.

ACH används för att systematiskt värdera hypoteser utifrån hur de stöds eller motsägs av tillgänglig information. I Hyppo används en anpassad variant av ACH där information kopplas till enskilda tillstånd i det morfologiska fältet istället för till hela hypoteser. Eftersom en matris med m dimensioner, med vardera n tillstånd, ger upphov till n^m möjliga hypoteser underlättas hanteringen av stora hypotesmängder väsentligt av att värderingen sker mot enskilda tillstånd.

Det första av de två användartesterna genomfördes med analytiker på FOI vid tre tillfällen där två analytiker deltog vid varje tillfälle. Det andra användartestet genomfördes med tio analytiker från Försvarsmakten som en avslutande del under en analysworkshop.

Hyppo som helhet fick positiva omdömen av testdeltagarna. Deltagarna såg behov och nytta av denna typ av verktyg och tyckte att design och gränssnitt var

enkelt och tilltalande. Ett av de mest centrala utvecklingsbehoven som poängterades under samtliga genomföranden var att kunna använda osäkerhetsfaktorer i analysen, till exempel genom att värdera informationens tillförlitlighet och relevans.

Nyckelord: Nyckelord: Morfologisk analys, Analys av konkurrerande hypoteser, analytiska verktyg, datorstöd, hypotesprövning

Summary

Hyppo is an analysis tool developed by FOI to support the process of generating and evaluating a large set of hypotheses. This report discusses the approach to, and the outcome of, two user studies conducted in order to study usability of Hyppo, as well as to identify short term and long term requirements for future releases.

Hyppo combines two established analytical methods, *Morphological Analysis* and *Analysis of Competing Hypotheses* (ACH). Morphological analysis is often used to specify and analyze possible future developments for so called *wicked problems*. During the morphological analysis process a morphological chart, or matrix, is created. The matrix includes a number of dimensions and a set of states for each dimension. In Hyppo, the matrix is used to generate hypotheses for the actual analysis problem. A hypothesis in Hyppo is composed of one state from each of the selected dimensions. As a simple example, we can consider a model for a forensic analysis with the dimensions *culprit*, *weapon* and *motive*. A specific hypothesis for this model could be “the culprit *Lee Harvey Oswald* with the weapon *Colt Holster Pistol* and the motive *political terrorism*”.

ACH is a methodology for systematic assessment of a set of hypotheses, based on how the available information supports or contradicts each hypothesis in the set. In Hyppo, an adjustment of standard ACH is introduced, where information is connected to specific states in the morphological chart, rather than to entire hypotheses. Since a chart consisting of $m \times n$ states will generate a set of n^m possible hypotheses, the analysis is facilitated by introducing this implicit procedure for hypothesis assessment.

Participants of the first of the two user studies were analysts at FOI. The study was conducted during three workshops, with two participants on each occasion. The second user study was carried out as part of a more extensive analysis workshop with ten participants from the Swedish Armed Forces.

A key future requirement identified was the ability to accommodate uncertainties in the analysis, for instance by means of expressing the accuracy, reliability and relevance of information introduced into the analysis. This was stressed by all participants. The overall testimonial was however favorable; the participants saw several benefits with this type of software and found the user interface appealing in its simplicity.

Keywords: Morphological Analysis, Analysis of Competing Hypotheses, analytical software tools, hypothesis testing

Innehållsförteckning

1	Inledning	9
1.1	Projektet	10
2	Beskrivning av analytiska metoder och verktyg	12
2.1	Morfologisk analys.....	12
2.2	Analys av konkurrerande hypoteser	13
2.3	Analytiska verktyg	14
3	Metod för utvärdering av Hyppo	16
3.1	Beskrivning av Hyppo.....	16
3.2	Scenarier	19
3.3	Förberedelser	21
3.4	Deltagare	22
3.5	Procedur	23
4	Resultat	26
4.1	Utvärdering av Hyppos stöd för analys	26
4.2	Utvärdering av Hyppos användargränssnitt	29
4.3	Övriga idéer	30
5	Diskussion	31
6	Slutsats	34
7	Referenser	35
	Bilaga 1. Utvecklingsbehov	37

1 Inledning

I denna rapport redovisar vi resultatet av användartest som genomförts under våren 2014 inom ramen för det FMV-finansierade projektet Multihypoteshantering, vars syfte är att undersöka om datorverktyget Hyppo, och det koncept för hypoteshantering som Hyppo stödjer, är användbart för att öka situationsförståelsen hos analytiker och beslutsfattare inom Försvarsmakten.

I underrättelsesammanhang genomförs analyser i syfte att till exempel förstå orsakssamband, diagnosticera problem och förutse möjliga framtida utvecklingar. I den här typen av analyser finns det ofta en relativt stor mängd alternativa hypoteser som i olika grad stöds eller motsägs av bevisunderlaget. Vidare är bevisunderlaget vanligtvis både ofullständigt och osäkert, och det gör att det i många analysproblem inte är möjligt att resonera sig fram till en entydig lösning.

Hypotesprövning är en viktig del av analytikers arbete. Hypotesprövning består av flera relaterade processer för att:

1. Utforma hypotesrum med rimliga hypoteser för givna problem,
2. Söka efter och formulera evidens och antaganden som kan användas till att värdera hypoteserna,
3. Pröva vilken eller vilka hypoteser som är rimligast.

Ett vanligt sätt att angripa denna typ av komplexa analysproblem är att snabbt bilda sig en uppfattning om vilken eller vilka hypoteser som är mest sannolika, och att sedan använda resten av bevisunderlaget för att bekräfta den uppfattningen. Nackdelen med denna strategi är att kvalitén på analysresultatet kan försämrats eftersom andra alternativ än den initiala uppfattningen aldrig prövas ordentligt.

För att förbättra kvalitén på analysen i problem med många alternativa hypoteser och ett omfattande och osäkert bevisunderlag utvecklar FOI ett datorstöd, kallat Hyppo, som baseras på de två etablerade analytiska metoderna *Morfologisk analys* och *Analys av konkurrerande hypoteser* (eng. *Analysis of Competing Hypothesis*, ACH). Grundidén med verktyget är att analytikern inte ska behöva begränsa analysen till en eller ett par hypoteser p.g.a. svårigheten att hantera mycket information manuellt.

Hypotesprövning är som många analytiska metoder en iterativ process som integrerar övergripande mål, tolkning, värdering och informationssökning (Pirolli & Card, 2005). Hyppo baseras på befintliga analytiska metoder för att genomföra processerna och skiljer sig från andra analysverktyg som antingen baseras på en metod eller kräver en helt ny analysmetod.

1.1 Projektet

Prototypverktyget Hyppo har i huvudsak utvecklats inom ramen för det pågående FMV-finansierade projektet Verktyg för informationshantering och analys (Gustavi, Johansson, Mårtenson & Svenson, 2011; Gustavi, Karasalo & Mårtenson, 2013). Viss vidareutveckling har också skett inom ramen för EU/FP7-projekten Contain¹ och Support². I det nu pågående FMV-finansierade projektet Multihypoteshantering ligger fokus på utvärderingar av Hyppos funktionalitet.

Syftet med projektet som helhet är att undersöka om stödverktyget Hyppo, och det koncept för hypoteshantering som Hyppo stödjer, är användbart för att öka situationsförståelsen hos analytiker och beslutsfattare inom Försvarmakten, samt att förbereda för införande av ett datorverktyg baserat på Hyppo i Försvarmaktens system för underrättelseanalys och informationshantering.

För att uppnå syftet har projektet följande mål:

- Ta fram en stabil grundplattform som i framtida projekt kan anpassas/byggas ut för att passa olika behov inom bland annat Försvarmakten.
- Identifiera, och i mån av tid implementera, mer avancerad funktionalitet som är intressant för Försvarmakten.
- Utvärdera verktygets potential inom Försvarmakten genom användarförsök.
- Presentera verktyget för lämpliga personer inom bl.a. Försvarmakten.

Projektet är uppdelat i tre arbetspaket:

- AP1: Genomförande av användartest med analytiker från FOI. Fokus på utvärdering av gränssnitt och grundläggande funktionalitet.
- AP2: Förstudie om kort- och långsiktiga vidareutvecklingsbehov för Hyppo samt, i mån av tid, implementering av ny funktionalitet i Hyppo inför användartest i AP3.
- AP3: Genomförande av användartest med analytiker från Försvarmakten. Fokus på utvärdering av analysstöd och arbetsprocesser.

I denna rapport redovisas resultatet av den första omgången användartest som genomförts under våren 2014 inom ramen för AP1. Det huvudsakliga syftet med

¹ <http://www.containproject.com/> [2014-06-19]

² <http://www.supportproject.info/> [2014-06-19]

den här redovisade användarstudien har alltså varit att studera verktygets gränssnitt och grundläggande funktionalitet, samt att identifiera utvecklingsbehov för kommande versioner, specifikt inför genomförande av kommande användarförsök i AP3, men även på längre sikt.

I FOI:s projektgrupp har Tove Gustavi, Robert Forsgren, Maja Karasalo, Peter Svenmarck, Magdalena Granåsen och Johanna von Feilitzen (examensarbetare från Uppsala Universitet) ingått.

2 Beskrivning av analytiska metoder och verktyg

Det inom projektet utvecklade prototypverktyget Hyppo bygger på de två analytiska metoderna *Morfologisk analys* och *Analys av konkurrerande hypoteser (ACH)*.

Morfologisk analys används vanligen för att konkretisera möjliga hypoteser eller scenarier i så kallade vildvuxna problem (Stenström, 2013). ACH är i sin tur utvecklad för att stödja analytikern vid hypotesvärdering. Metoderna beskrivs översiktligt i kapitel 2.1 och 2.2. För mer utförliga metodbeskrivningar hänvisas läsaren till (Gustavi, Johansson, Johansson, Jändel, Kaati, Mårtenson, & Oskarsson, 2011).

Genom att kombinera morfologisk analys och ACH stödjer Hyppo:

- Utformning av ett morfologiskt fält som ger strukturen åt analysproblemet och kan användas för hypotesgenerering
- Registrering och bearbetning av evidens
- Värdering av hur evidensen stödjer eller motsäger de tillstånd som tillsammans utgör en hypotes
- Överblick av hur mycket stöd hypoteserna har av evidensen

Dessa funktioner underlättar systematisk prövning av många hypoteser från ett omfattande bevisunderlag. Hyppo syftar därför till att ge analytikern bättre förutsättningar att bedöma vilka hypoteser som är rimliga, förstå hur osäkerheter påverkar bedömningen och identifiera kritisk information för fortsatta analyser.

2.1 Morfologisk analys

Metoden *morfologisk analys* är särskilt utvecklad för att formulera, strukturera och studera komplexa problem (Stenström, 2013). Morfologisk analys utvecklades ursprungligen för att studera problem inom astronomi och men generaliserades sedan till mångdimensionella problem som är svåra att kvantifiera (Zwicky, 1969; Ritchey, 2006). "Morfo" härstammar från grekiskans ord för "form", och morfologi handlar om också om formlära och hur delar skapar en gemensam helhet. FOI använder oftast Morfologisk analys i fokusgrupper för att utifrån en relevant frågeställning strukturera och formulera komplexa problem (Stenström, 2013).

Under den morfologiska analysen skapar analytikern en morfologisk modell som kan användas för att på ett systematiskt sätt generera tänkbara hypoteser eller scenarier för ett givet analysproblem. Den morfologiska modellen är utformad

som en matris (se exempel i Figur 1). Matrisen byggs upp av ett antal dimensioner som analytikern bedömer är viktiga för att beskriva en hypotes eller ett scenario, samt möjliga tillstånd för varje dimension. I en mordutredning kan till exempel en viktig dimension i en hypotes vara ”mordvapen”. De relaterade tillstånden är då olika vapen som kan ha använts, såsom ”pistol” eller ”kniv”. Dimensionerna ska gärna vara oberoende av varandra och kan av pedagogiska skäl sorteras i prioritetsordning från vänster till höger i matrisen, beroende på vad som är mest betydelsefullt för den aktuella analysen. Kombinationer av tillstånd från de olika dimensionerna representerar scenarier som i Hyppo benämns som hypoteser. Vissa kombinationer av tillstånd kan vara logiskt eller empiriskt omöjliga, till exempel genom att de förutsätter att en person befunnit sig på två olika platser samtidigt, och sådana kombinationer kan då exkluderas från den fortsatta analysen.

Gärningsman	Mordvapen	Brottsplats
Anna	Kniv	Hotellrummet
Bertil	Pistol	Hotellspat
Cecilia		Garaget
David		

Figur 1. Exempel på ett morfologiskt fält. Gärningsman, Mordvapen och Brotsplats utgör dimensioner. Varje dimension har ett antal möjliga tillstånd.

Kombinationer av tillstånd används i Hyppo för att generera ett representativt urval av alla mer eller mindre tänkbara hypoteser för ett visst analysproblem. Fördelen med att kombinatoriskt generera hypoteser utifrån den morfologiska modellen, istället för att låta analytikern direkt och utan mellansteg ange vilka hypoteser som ska analyseras, är att analytikern konfronteras med hypoteser som vid en första anblick kan ses som osannolika, men som vid närmare betraktande ändå kan visa sig vara rimliga givet tillgängliga evidens.

I denna rapport används begreppet *morfologiskt fält* istället för *morfologisk modell*, eftersom Hyppo i nuläget inte fullt ut stödjer arbetsprocessen i traditionell morfologisk analys.

2.2 Analys av konkurrerande hypoteser

Metoden *Analys av konkurrerande hypoteser* (ACH) har utvecklats av den tidigare CIA-anställda analytikern Richard Heuer. Idén bakom ACH är att en strukturerad arbetsprocess ska minska risken för att hypotesprövning påverkas av

analytikerns initiala uppfattning (eng. *confirmation bias*). Genom att värderingen av alla evidens och antaganden görs explicit blir det också lättare att motivera varför den valda hypotesen är rimligast. ACH består av följande åtta steg (Heuer, 1999; jmf. Jones, 1998):

1. Identifiera alla hypoteser,
2. Lista alla relevanta evidens,
3. Skapa en tabell för alla hypoteser (kolumnvis) och evidens (radvis). Värdera hur varje evidens påverkar trovärdigheten för var och en av hypoteserna, och för in resultatet i tabellen,
4. Förfina tabellen genom att förtydliga hypoteser, ta bort evidens som inte bidrar till hypotesprövningen och lägga till nya informativa evidens,
5. Jämför hypoteserna,
6. Gå igenom resultatet,
7. Identifiera uppföljningsindikatorer, dvs händelseförlopp eller ny information som kan påverka analysens giltighet.
8. Rapportera resultatet.

Värt att notera är att man i ACH skiljer generering av information, antaganden och argument från själva hypotesprövningen. Först listas all relevant information (evidens), därefter värderas alla evidens mot alla hypoteser och först därefter jämfört hypoteserna mot varandra. På så sätt är tanken att undvika att i förtid låsa på en viss hypotes. Studier visar hur ACH minskar risken för confirmation bias hos mindre erfarna analytiker (Lehner, Adelman, Cheikes, & Brown, 2008).

I ACH, liksom i vetenskaplig hypotesprövning, betonas vikten av evidens som motsäger en hypotes eftersom de kan användas för att utesluta hypoteser och därmed minska omfattningen av det analyserade problemet. För ACH rekommenderas att max 7 hypoteser hanteras samtidigt, annars blir analysen svåröversiktlig (Heuer, 1999). Att gruppera hypoteser genom att generalisera dem är ett sätt att hantera detta. I Hyppo finns ett analysstöd som gör det möjligt att hantera ett betydligt större antal hypoteser i analysprocessen än med traditionell ACH. Det innebär att man kan hantera detaljerade hypoteser och slipper utesluta hypoteser innan de på ett trovärdigt och pålitligt sätt blivit motbevisade.

2.3 Analytiska verktyg

I denna sektion presenteras verktyg som innehåller analysstöd på liknande sätt som Hyppo Dessa verktyg innehåller stöd för att strukturera och genomföra analyser utifrån morfologisk analys eller evidensbaserad hypotesprövning.

Verktöget Casper (Computer Aided Scenario and Problem Evaluation Routine) är särskilt utvecklat för att genomföra morfologiska analyser (Stenström, 2013). Casper stödjer bland annat utformningen av morfologiska fält och exkluderingen av orimliga tillståndskombinationer. Olika möjliga tillståndskombinationer bildar scenarier. Casper har däremot inget stöd för evidensvärdering eller värdering av de olika scenariernas giltighet.

Verktöget PARC ACH är särskilt utvecklat för att genomföra analys av konkurrerande hypoteser (PARC, 2010). I PARC ACH anges evidens och hypoteser i en tabell. PARC ACH har även enklare stöd för att skatta trovärdighet och relevans för varje evidens i en tregradig skala (låg, medel och hög). Trovärdighet är analytikerns egen sammanvägning av sakriktighet och källans tillförlitlighet medan relevans relaterar till hur viktigt evidenset är för analysen. Dessa skattningar översätts till numeriska värden och analytikern kan sedan välja om dessa ska viktas in i analysen. Ett evidens som har lägre skattningar på relevans eller trovärdighet får alltså mindre betydelse i analysen. PARC ACH har inte något stöd för att bygga upp morfologiska fält.

Ett exempel på ett mer avancerat analysverktyg för underrättelseanalys är Impactorium. Liksom Hyppo är Impactorium avsett att stödja analytiker som jobbar med hypoteser, men angreppssättet är helt annorlunda. Impactorium har utvecklats för att automatiskt värdera hur evidens stödjer ett antal fördefinierade hypoteser (Forsgren, Kaati, Mårtenson, Svenson, & Tjörnhammar, 2008). Ett annat verktyg är MHSA (Multiple Hypothesis Situation Analysis) som ger stöd för att hitta möjliga hypoteser och utifrån evidens identifiera de viktigaste hypoteserna för fortsatta analyser (Roy & Guyard, 2010). MHSA fokuserar på kvantitativa analyser snarare än analytikerns förståelse av problemet. TH!NK Suite liknar Hyppo på så sätt att det kombinerar morfologisk analys och ACH (Globalytica, 2014), men eftersom övergången till ACH är ett separat steg så innebär det i praktiken att många hypoteser måste väljas bort redan innan evidensvärderingen.

3 Metod för utvärdering av Hyppo

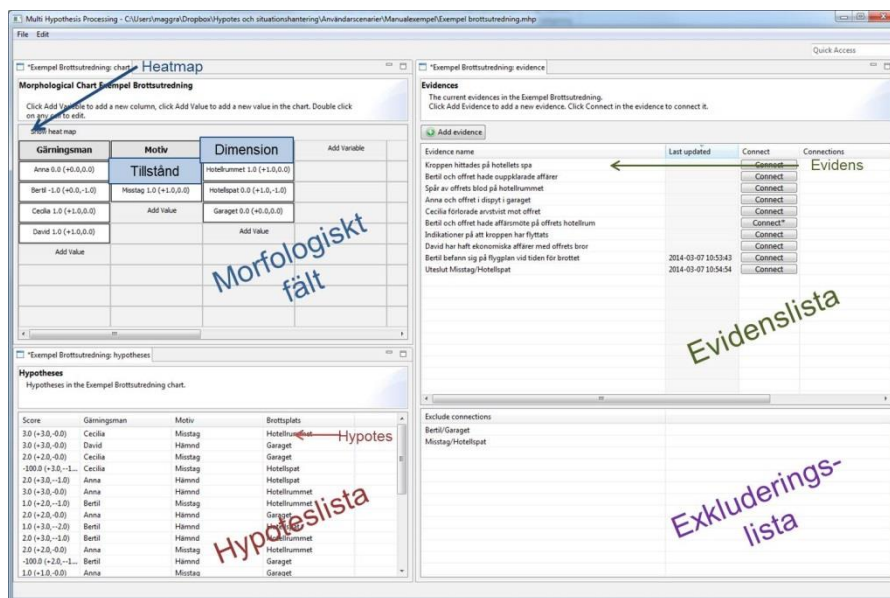
Projektgruppen genomförde två användartester under våren 2014 i syfte att utvärdera Hyppos gränssnitt och grundläggande funktionalitet, samt att identifiera utvecklingsbehov för kommande versioner. Användartest 1 genomfördes med analytiker från FOI. Användartest 2 genomfördes med analytiker från Försvarmakten som en avslutande del under en workshop om analytiska metoder.

Föreliggande kapitel beskriver verktyget och analysprocessen för Hyppo, de scenarier som användes under användartesten, deltagare och metod för genomförande och datainsamling.

3.1 Beskrivning av Hyppo

Hyppo är utvecklat för att förbättra och underlätta analysen av många alternativa hypoteser utifrån ett omfattande och osäkert bevisunderlag. Hyppo är baserat på de analytiska metoderna Morfologisk analys och ACH. Gränssnittet består av fyra fönster som i tur och ordning innehåller *morfologiskt fält*, *evidenslista*, *hypoteslista* och *exkluderingslista* (se Figur 2). I det morfologiska fältet definieras och redigeras dimensioner och möjliga tillstånd. I evidenslistan finns de evidens som har registrerats under analysen. Hypoteslistan visar alla hypoteser som genererats utifrån det morfologiska fältet. Exkluderingslistan visar de hypoteser (dvs. tillståndskombinationer) som analytikern valt att utesluta ur analysen.

Hypoteslistan kan sorteras baserat på hur evidens stödjer, motsäger eller exkluderar de olika hypoteserna. När en hypotes väljs i hypoteslistan så markeras motsvarande tillstånd i det morfologiska fältet. Det morfologiska fältet visar också det totala antalet evidens som stödjer och motsäger varje tillstånd. En *heat map* (se exempel i Figur 3) kan överlagras på det morfologiska fältet. Genom en färgskala från grönt till rött som ges en övergripande uppfattning om vilket stöd enskilda tillstånd har gentemot evidensen.



Figur 2. Hypos gränssnitt uppdelat i morfologiskt fält, evidenslista, hypoteslista och exkluderingslista

Gärningsman	Motiv	Orsak	Förlopp
Amerikansk militär 1.0 (+1.0,0.0)	Vildskott 0.0 (+1.0,-1.0)	Värmedskande robot -1.0 (+2.0,-3.0)	Sprängmedelsexplosion 1.0 (+2.0,-1.0)
Terrorist 1.0 (+1.0,0.0)	Mänskliga faktorn -2.0 (+0.0,-2.0)	Ballistisk robot 1.0 (+2.0,-1.0)	Brändeexplosion 0.0 (+1.0,-1.0)
Annan kriminell 0.0 (+0.0,0.0)	Politiskt motiv 1.0 (+1.0,0.0)	Tekniskt fel 0.0 (+0.0,0.0)	Annat primärförlopp än explosion -1.0 (+0.0,-1.0)
Flygplansövervakaren -2.0 (+0.0,-2.0)	Egen vinning 0.0 (+0.0,0.0)	Sabotage 0.0 (+0.0,0.0)	Add Value
Underhållstjänsten -2.0 (+0.0,-2.0)	Add Value	Robot med bildmålare 1.0 (+2.0,-1.0)	
Forsvarsindustri 0.0 (+0.0,0.0)		Sprängmedel ombord 0.0 (+1.0,-1.0)	
Add Value		Flygbomb 0.0 (+1.0,-1.0)	Add Value

Figur 3. Heatmap överlagrad på det morfologiska fältet. Tillstånd markerade med grönt har fler stödande än motsägande evidens. Tillstånd markerade med rött har fler motsägande evidens.

I Hyppo genereras hypoteser utifrån det morfologiska fältets tillstånd, där varje hypotes består av ett tillstånd från varje dimension i det morfologiska fältet. ACH används för att värdera de enskilda tillstånden i det morfologiska fältet. Programmet räknar sedan ut hur mycket stöd var och en av hypoteserna har, utifrån hur mycket stöd tillstånden som ingår i respektive hypotes har fått i ACH-analysen. Eftersom en tabell med $n \times m$ tillstånd ger upphov till n^m möjliga

hypoteser underlättas hanteringen av stora hypotesmängder väsentligt av att värdering sker mot enskilda tillstånd istället för kombinerade hypoteser.

En analys med Hyppo består av följande steg:

Steg 1) Iterativt registrera och bearbeta evidens och utforma det morfologiska fältet: Utifrån en problemformulering skapas ett morfologiskt fält med de dimensioner som anses lämpliga givet problemets natur och tillgänglig information. För ostrukturerade problem är det ofta nödvändigt att iterera problemformuleringen och evidensen flera gånger för att bilda sig en uppfattning om hur man på lämpligt sätt ska utforma sitt morfologiska fält. Analytikern kan behöva pröva olika sätt att strukturera det morfologiska fältet. I den iterativa processen bearbetas också evidensen så att de kan användas för att värdera tillstånd i det morfologiska fältet, till exempel kan ett evidens behöva delas upp i flera evidens eller så kan man behöva formulera ett nytt evidens som ersätter flera befintliga evidens med överlappande informationsinnehåll. Det är viktigt att dimensionerna i det morfologiska fältet är meningsfulla, entydiga och oberoende, samt har lämplig abstraktionsnivå. Dimensionerna kan också ordnas logiskt eller pedagogiskt för att underlätta hypotesvärdering.

Steg 2) Evidensvärdering: När det morfologiska fältet till slut avspeglar strukturen på analysproblemet värderas alla evidensen efter hur de stödjer eller motsäger tillstånden i det morfologiska fältet. Värderingen sker genom att markera ett evidens och för detta evidens ange de tillstånd som stödjer respektive motsäger evidenset. Ovanför det morfologiska fältet finns knappar för positiv (stödjande), negativ (motsägende) eller neutral (ingen värdering) koppling. Samma evidens kan stödja vissa tillstånd och motsäga andra. I det morfologiska fältet visas antalet stödjande och motsägende evidens i en parentes vid varje tillstånd.

Steg 3) Exkludera orimliga hypoteser: Eftersom hypoteserna automatiskt genereras i form av tillståndskombinationer från det morfologiska fältet (ett tillstånd per dimension) uppstår ofta hypoteser som innehåller orimliga tillståndskombinationer. Kombinationerna kan vara logiskt uteslutande eller av annan orsak bedömas vara orimliga. Hypoteser som innehåller dessa kombinationer kan då exkluderas. I dagsläget behöver denna exkludering ske inom ramen för ett evidens. Finns det inte något evidens som stödjer en viss uteslutning så behöver ett nytt evidens skapats. När ett evidens valts kan tillstånd som tillhör två olika dimensioner i det morfologiska fältet exkluderas från varandra med hjälp av en knapp ovanför det morfologiska fältet. Orimliga tillståndskombinationer kan uppstå mellan två dimensioner men också omfatta fler dimensioner.

Steg 4) Hypotesbedömning: När alla evidens är värderade och orimliga tillståndskombinationer exkluderade kan en bedömning göras av vilka hypoteser som är mest rimliga. Vid bedömningen är motsägende evidens generellt mer

intressanta än stödjande evidens eftersom de i vissa fall kan användas för att utesluta hypoteser. I brist på tydligt motsägande evidens så kan kombinationer av stödjande evidens ge belägg för att vissa hypoteser är rimligare än andra. I den nuvarande versionen av Hyppo beräknas hypotesvärderingen utifrån skillnaden mellan antalet stödjande och motsägande evidens. Hyppos funktion *Heatmap* ger en visuell bild av vilka tillstånd i det morfologiska fältet som har flest positiva respektive negativa evidens kopplade till sig.

Steg 5) Rapportera: Eftersom många analysproblem saknar en entydig lösning så måste rapporteringen efter genomförd analys beskriva den bild som evidensen ger av det morfologiska fältet. En sammanvägning av stödjande och motsägande evidens, osäkerheter i evidensen och vilken kritisk information som kan påverka analysresultatet ger en uppfattning om hypotesernas rimlighet. Rapporteringen bör sammanfatta detta resonemang och identifiera väsentliga osäkerheter.

Ett processdiagram motsvarande Steg 1 och 2 kan ses i Figur 4.

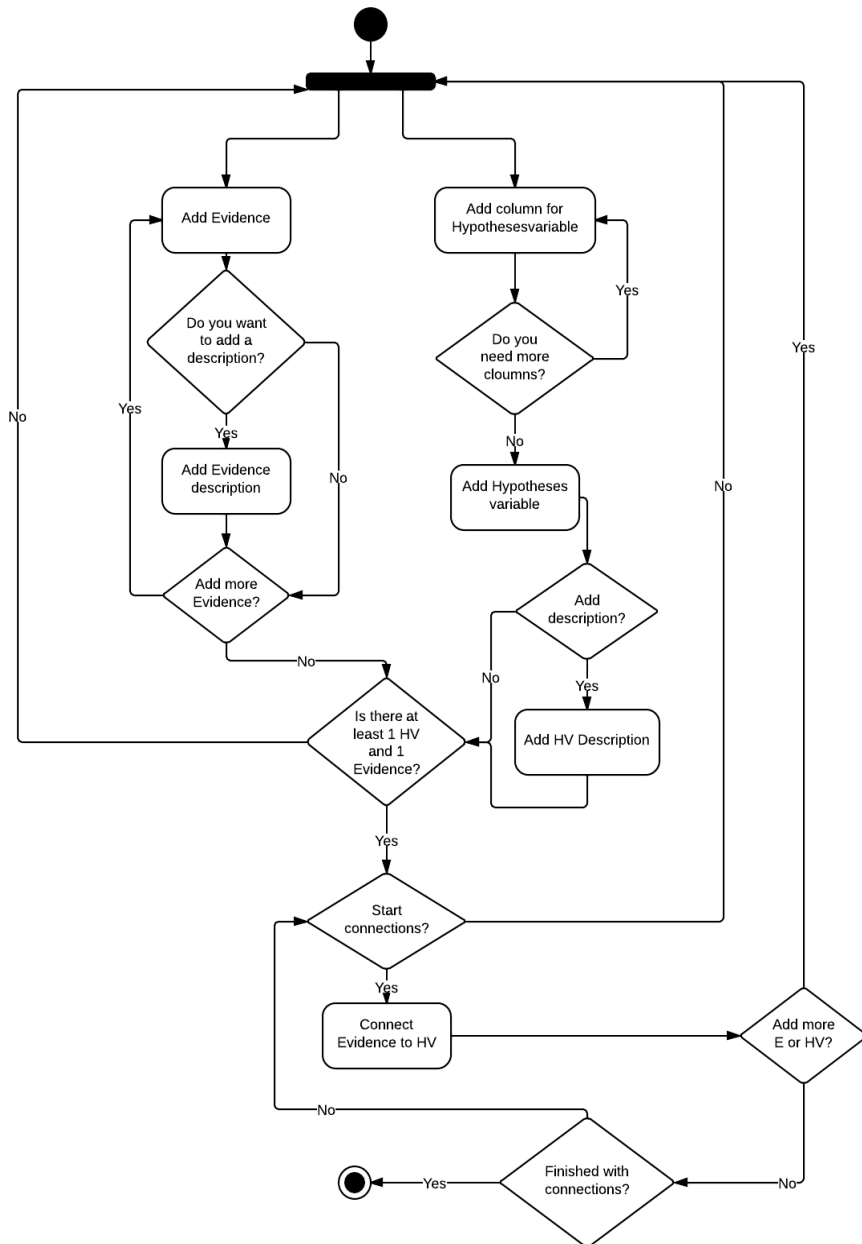
3.2 Scenarier

Två olika scenarier användes i användartesterna – ”Flygkrasch TWA 800” och ”ALS-PD hos Guams ursprungsbefolkning”. Ursprungstanken var att det enklare TWA800-scenariot skulle utgöra träningsscenario och det mer omfattande och komplexa Guam-scenariot skulle utgöra testscenario vid varje genomförande. Under första försöket visade det sig emellertid att även TWA800-scenariot var tillräckligt omfattande för den avsatta tiden och samtidigt tillräckligt komplext för att testa Hyppos funktionalitet och identifiera utvecklingsbehov.

Testledningen valde därför använda Guam-scenariot enbart vid ett av tre tillfällen under användartest 1, nämligen vid det tillfälle då försöksdeltagarna var experter inom morfologisk analys. Övriga deltagare i användartest 1 samt användartest 2 analyserade TWA800-scenariot.

Scenariot ”Flygkrasch TWA 800” är baserat på en verklig olycka som inträffade 1996 när ett passagerarflygplan på väg från New York till Paris kraschade i Atlanten utanför New York endast minuter efter start. Samtliga passagerare omkom och det finns en mängd olika teorier om vad som orsakade olyckan. Några exempel på möjliga olycksorsaker är en bomb ombord på planet, träff av en luftvärnsrobot eller annan missil eller bränsleexplosion till följd av tekniskt fel. Försöksdeltagarnas uppgift var att identifiera den mest troliga orsaken till kraschen utifrån tillgänglig information.

Scenariot ”ALS-PD hos Guams ursprungsbefolkning” är även det baserat på verkliga händelser. ALS-PD är en ovanlig sjukdom som nästan enbart drabbar Chamorros, ursprungsbefolkningen på Guam. ALS-PD påminner om Amytrofisk lateralskleros (ALS), Alzheimer och Parkinsons sjukdom. Några exempel på möjliga orsaker till sjukdomen relaterar till beredning av mjöl från cykadfrön,

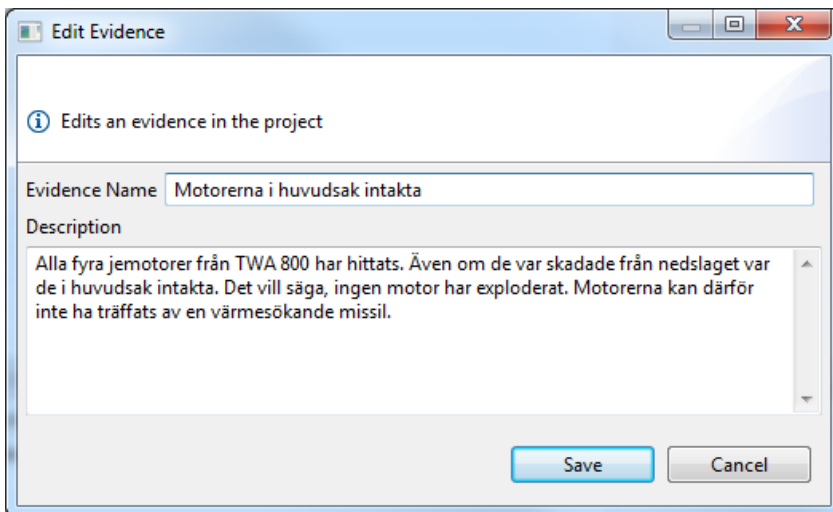


Figur 4. Arbetsordning för steg 1 och 2 i Hyppos analysprocess. Ur von Feilitzen (2014).

intag av bröd baserat på detta mjöl, användning av cykadblad för medicinering och läkning av sår, maträtter på fladdermöss eller genetiska orsaker. Försöksdeltagarnas uppgift var att identifiera den eller de mest troliga orsakerna till sjukdomen.

3.3 Förberedelser

För att spara tid matade testledarna huvuddelen av scenarioinformationen som evidens i Hyppos evidenslista före användartestet (Figur 5). Bara ett par evidens utelämnades för att ge deltagarna möjlighet att själva testa inmatningen av evidens. Testledarna skapade även ett utkast till morfologiskt fält i förväg för att spara tid. Bara några tillstånd utelämnades för att ge deltagarna möjlighet att pröva funktionerna för utformning av det morfologiska fältet. Utöver att komplettera evidenslista och morfologiskt fält med information som saknades så uppmuntrades deltagarna att ändra det initiala morfologiska fältet om de tyckte att en annan indelning var lämpligare. När användartesten började fanns inga kopplingar mellan evidenslista och morfologiskt fält och inga kombinationer av tillstånd hade exkluderats.



Figur 5. Exempel på inmatat evidens i Hyppo (Scenario TWA 800).

3.4 Deltagare

3.4.1 Deltagare i användartest 1

Totalt deltog sex analytiker i användartest 1, två vid varje testtillfälle. Samtliga analytiker var anställda vid FOI och hade titeln analytiker/forskare, forskningsledare eller forskningsingenjör, de flesta med gedigen erfarenhet av analytiskt arbete. Två deltagare hade specifikt valts ut för sin erfarenhet inom morfologisk analys och de arbetade tillsammans i ett par. Det här paret analyserade Guam-scenariot eftersom detta är ett mer svårstrukturerat problem och därmed ger ett större fokus på den morfologiska delen av Hyppo. De två andra paren analyserade TWA800-scenariot.

Bakgrundsenkäten visade en spridning på 2-33 års erfarenhet av analytiskt arbete, med ett genomsnitt på 20 års erfarenhet (median 20 år). Några deltagare räknade även in sina universitetsstudier som analytiskt arbete. Analyserfarenhet beskrevs i termer av militär underrättelseanalys, militär operationsanalys, säkerhetspolitiska analyser, analyser för olika civila myndigheters praktiska problem, och utredningsarbete. Deltagarna beskrev också vad de ansåg kännetecknar analytiskt arbete. Sammanfattningsvis nämndes strävan efter objektivitet, förmåga att urskilja relevant information, förmåga att se mönster, helhet och samband, systematisk analys och syntes, alternativgenerering och -värdering, problemformulering och -avgränsning, aggregering av data och kritisk granskning.

Samtliga deltagare hade tidigare kommit i kontakt med morfologisk analys, men ingen hade arbetat med ACH. Samtliga deltagare hade också någon erfarenhet av tekniska beslutsstödsverktyg. Verktyg för stöd till analys som nämndes var Casper (fem deltagare) (Stenström, 2013), Officepaket (främst Excel, tre deltagare), MindManager³ (en deltagare), Konceptkartor⁴ (en deltagare), Expert Choice⁵ (en deltagare), Gamma (en deltagare), Hugin Expert⁶ (en deltagare) och Optima (en deltagare) (Carlsen & Dreborg, 2008). Deltagarna skattade sin tekniska förmåga att hantera IT-system till i genomsnitt 3,7 på en femgradig skala.

I intervjuer som genomfördes av en examensarbetare i projektet framkom att försöksdeltagarnas arbetssituationer varierar, både sinsemellan och över tid; till exempel så uppgav endast två av sex försöksdeltagare att de regelbundet arbetar i grupp med analysarbete och fyra av försöksdeltagarna uppgav att deras fysiska arbetsplats varierar beroende på uppdrag. Gemensamt för testdeltagarna var att

³ <http://mindmanager.se/> [2014-06-27]

⁴ Till exempel IHMC CMapTools, <http://cmap.ihmc.us/> [2014-06-27]

⁵ <http://expertchoice.com/> [2014-06-27]

⁶ <http://www.hugin.com/> [2014-06-27]

de alla anmält sig frivilligt som deltagare i användarförsöken p.g.a. ett personligt intresse av att se nya analytiska verktyg. Det kan inte uteslutas att detta faktum kan ha påverkat resultatet av utvärderingen i positiv riktning. Generellt ansåg försöksdeltagarna att de i nuläget inte har problem med att göra sina analyser i huvudet, men de uppgav trots det att de skulle behöva verktyg som gav stöd för beräkningar och som kunde ge nya uppslag. I nuläget är det vanligare med kvalitativa än kvantitativa analyser och det framkom under intervjuerna att det finns ett intresse av att göra fler kvantitativa analyser, men att det då behöver finnas tekniska verktyg som stödjer detta. Fem av sex försöksdeltagare uppgav att de ”ofta och mycket” samarbetar med andra personer och såg därför fördelar med att använda sig av samma modeller när de dokumenterar, rapporterar och arbetar. En mer detaljerad sammanställning av intervjuerna finns att läsa i rapporten ”Hur analytiska beslutsstödsystem kan stödja underrättelsetjänstens arbete inom den svenska Försvarsmakten” (von Feilitzen, 2014).

3.4.2 Deltagare i användartest 2

I användartest 2 deltog tio analytiker ur Försvarsmakten. Två av deltagarna arbetade med utbildningsverksamhet för analytiker medan övriga var specialister inom olika sakområden. Några analytiker hade militär bakgrund medan andra var civilanställda. Samtliga deltagare hade provat på de båda metoderna morfologisk analys och ACH var för sig, samt provat på att arbeta i verktyget PARC ACH. Deltagarna hade även arbetat med TWA800-scenariot under den tidigare delen av workshopen. Eftersom ingen bakgrundsenkät genomfördes finns inte någon ytterligare information om teknisk kompetens eller deltagarnas syn på och erfarenhet av analytiskt arbete.

3.5 Procedur

Användartest 1 genomfördes under tre tre-timmarsspass. Vid varje tillfälle deltog två försöksdeltagare, två-tre projektdeltagare för utbildning, handledning och datainsamling samt en examensarbetare som samlade in data utifrån sina frågeställningar. Deltagarna för användartest 1 fyllde i en *bakgrundsenkät* enskilt före användartestet. Bakgrundsenkäten innehöll frågor om deltagarnas syn på vad som kännetecknar analytiskt arbete, deras erfarenheter av ACH respektive morfologisk analys, erfarenheter av att arbeta med tekniska beslutsstödsverktyg för analys, samt skattning av sin tekniska förmåga att hantera IT-system (1-5).

Användartest 1 inleddes med att examensarbetaren genomförde en 20-minuters intervju enskilt med varje deltagare i syfte att få en bättre uppfattning om hur deras arbetssituation såg ut mer generellt. Intervjuerna behandlade bland annat ämnen som arbetsrutiner, användande av analytiska verktyg (ej begränsat till morfologisk analys och ACH), inställning till datorstödda metoder, former för samarbete med andra analytiker och önskemål om nya stödfunktioner.

Under tiden fick den andra deltagaren i paret utbildning på Hyppo och arbetsmetoderna. Därefter arbetade försöksdeltagarna under ca 90 minuter tillsammans med att lösa analysuppgiften utifrån det förinmatade scenariot. Två av paren analyserade TWA 800-scenariot medan ett av paren analyserade Guamscenariot. Parets uppgifter var att:

- Komplettera evidenslistan med den information som inte matats in.
- Komplettera och modifiera det morfologiska fältet.
- Värdera evidens gentemot det morfologiska fältet genom att göra kopplingar mellan dessa två.
- Exkludera orimliga kombinationer av tillstånd i det morfologiska fältet (utifrån starkt motsägende evidens eller logisk orimlighet).
- Göra en värdering av de tre mest troliga hypoteserna utifrån analysen. Deltagarna kunde då välja de hypoteser som Hyppo förespråkade (högst poäng) eller någon annan hypotes i hypoteslistan.
- Löpande kommentera behov av funktionalitet, brister i användarvänlighet och andra synpunkter på Hyppo.

Under lösandet av analysuppgiften antecknade testledarna löpande de utvecklingsbehov som uppkom och frågade vid behov om förtydliganden. Utvecklingsbehoven dokumenterades i en tabell samt i fritext.

Efter lösandet av analysuppgiften fick deltagarna i användartest 1 enskilt fylla i en *efterenkät* (ett par valde att göra detta tillsammans på grund av tidsbrist). Detta bedöms inte påverkat validiteten då det var tydligt att de var helt överens om svaren). Efterenkäten bestod till stor del av påståenden som deltagarna skulle ta ställning till på en femgradig skala från Instämmer inte alls till Instämmer helt. Det fanns också fritextfält att motivera sina val. Frågorna relaterade till

- Överensstämmelsen mellan analysverktygets resultat och deltagarnas egen känsla för mest trolig hypotes,
- I vilken utsträckning Hyppos olika funktioner var ett stöd för analysen (evidenslistan, viktningsfunktionen, det morfologiska fältet, exkluderingsfunktionen, Heatmap o.s.v.),
- Tillgänglig tid för lösande av analysuppgift och om analysproblemet var adekvat för uppgiften,
- Om användning av verktyget leder till en snabbare eller långsammare analys med sämre eller bättre kvalitet,
- Gränssnitt i form av hur enkelt Hyppos upplevdes att använda, gav en överblick över analysen och visualisering av den information som matats in.

Användartestet avslutades med en gemensam *diskussion* som en projektmedlem faciliterade utifrån förberedda teman som relaterade till verktygets nytta och utvecklingsbehov.

Användartest 2 genomfördes som ett avslutande moment under en tredagars workshop om analytiska metoder som inkluderade bland annat morfologisk analys och ACH. Deltagare under workshopen var analytiker vid Försvarsmakten. Efter en genomgång av Hyppo fick deltagarna under ca 30 minuter lösa det förinmatade TWA 800-scenariot i två grupper. Eftersom deltagarna hade kommit i kontakt med TWA 800-scenariot tidigare under workshopen visste de vilken information som scenariot innehöll. Därefter faciliterade en projektmedlem en diskussion med hela gruppen på samma sätt som under användartest 1. Inga enkäter delades ut under användartest 2 eftersom tiden inte medgav detta.

4 Resultat

Resultaten från användartesterna sammanställdes utifrån enkäter, observationer och diskussioner under och efter genomförandet av analysuppgiften. Enkäterna besvarades bara av deltagarna i användartest 1. Eftersom det var så få deltagare redovisas endast de enkätsvar där deltagarna var i stort sett eniga, dvs. där samtliga svar låg inom två skalsteg på den femgradiga skalan.

Deltagarna kommenterade utvecklingsbehov för Hyppo allt eftersom under analysen, vilket testledarna dokumenterade. Enkätsvaren från användartest 1 (skala 1-5) visade att deltagarna svarade positivt på om analysproblemet var representativt för typiska analyser (M=4,5) och om utbildningen på Hyppo var tillräcklig (M=4,5), men att de hade för lite tid för att lösa uppgiften (M=1,8). Tidsbristen bedöms inte nämnvärt ha påverkat användartestets syfte att pröva Hyppos funktionalitet och identifiera utvecklingsbehov. Även om deltagarna inte hann med att finslipa sina analyser och exkludera alla orimliga tillståndskombinationer så innebar den iterativa processen att alla testgrupper testade alla Hyppos funktioner. Utvecklingsbehoven som projektgruppen identifierade under användartesterna redovisas i Bilaga 1. I övrigt löste grupperna uppgiften utan friktioner, dvs. deltagarna förstod sin uppgift och utnyttjade Hyppos funktioner på avsett sätt.

Resultaten har strukturerats enligt utvärdering av Hyppos stöd för analys, dess gränssnitt och övriga idéer.

4.1 Utvärdering av Hyppos stöd för analys

Deltagarna var generellt positiva till Hyppos möjligheter. De ansåg att Hyppo ger förutsättningar för en mer strukturerad analys med bättre kvalitet, men förkortar inte tiden för analysen. En deltagare uttryckte att genom att alla möjliga hypoteser genereras så finns det möjlighet att bli uppmärksam på hypoteser som man annars inte skulle ha tänkt på. Resultaten presenteras med avseende på vilket stöd Hyppo gav för utformningen av det morfologiska fältet, evidensvärdering, tidseffektiv analys med bra kvalitet, samt dokumentation.

4.1.1 Morfologiskt fält och exkludering av tillståndskombinationer

Utformningen av ett morfologiskt fält med Hyppo prövades i första hand med det komplexa och svårstrukturerade Guam-scenariot som två experter inom morfologisk analys analyserade. Paret strukturerade om det förberedda morfologiska fältet på ett imponerande sätt, men tyvärr gjorde den nuvarande implementeringen av Hyppo att systemet blev långsamt och kraschade när det morfologiska fältet växte. Hyppos nuvarande begränsning är högst 5-7

dimensioner med 5 tillstånd per dimension, beroende på vilken dator som används. Guam-scenariot visade sig kräva ett ännu större morfologiskt fält.

Det morfologiskt erfarna paret ansåg att det finns behov av att kunna skapa undermodeller, vilket kan vara användbart för Guam-scenariot. De ansåg att modeller som har fler än ca 7 dimensioner blir svåra att överblicka och istället är det bättre att skapa en undermodell. Den överordnade modellen och undermodellerna måste också vara kopplade och det ska vara enkelt att växla mellan modellerna. Paret var också intresserat av att kunna skapa alternativa modeller som använde samma evidenslista.

Ett utvecklingsbehov som deltagarna diskuterade ofta var ett enklare sätt att exkludera tillståndskombinationer eftersom det i dagsläget upplevs som omständligt och tidskrävande. Tillvägagångssättet i den nuvarande versionen är för omständligt där ett särskilt evidens ibland skapas för att kunna göra och visa exkluderingar. Deltagarna ansåg också att exkluderingar kan ha olika karaktär genom att vissa tillståndskombinationer är logiskt omöjliga medan andra bara är mindre troliga. Omöjliga kombinationer kan uteslutas helt eller visas på en separat flik medan mindre troliga kombinationer kan finnas kvar som hypoteser men kopplas till motsägande evidens.

4.1.2 Evidenslista och evidensvärdering

De mest centrala synpunkterna på verktyget berörde evidenslistan. Många deltagare önskade en möjlighet att kunna värdera evidens mot kombinationer av tillstånd och inte bara mot de enskilda tillstånden. En deltagare menade till och med att evidens inte bör värderas mot tillstånd utan mot hela hypoteser, dvs. tillstånd från alla dimensionerna. För att kunna värdera evidens mot tillståndskombinationer krävs funktionalitet för att koppla ihop tillstånd.

Ett av paren menade att evidens kan stödja en kedja av händelser, vilket är svårt att ange i den nuvarande versionen.

Samtliga deltagare i båda användartesterna poängterade att det behövs stöd för att hantera osäkerhet rörande evidens och för att särskilja mellan evidens som är antaganden och evidens som baseras på observationer och iakttagelser. Ett förslag från några deltagare var att det skulle kunna göras i termer av *tillförlitlighet* och *sakriktighet* enligt den terminologi som används i Försvarsmaktens Underrättelsereglemente (Försvarsmakten, 2010). Några deltagare menade också att det skulle vara bra att kunna byta sådana värderingskategorier så att Hyppo även kan användas inom andra domäner. Deltagarna poängterade att det måste vara enkelt att förstå hur osäkerhetsbedömningar ska göras och skala måste vara tydlig så att inte analysen blir onödigt komplicerad. En deltagare föreslog mer studier av bevisvärdering och resonemang inom brottsutredningar för att få hjälp med osäkerhetshanteringen.

En del av deltagarna under användartest 2 som fått prova att använda verktyget PARC ACH ansåg att dess kategorisering för osäkerheter i evidens (trovärdighet och relevans anges för varje evidens i låg, medel och hög) var enkel att använda, men andra ansåg att underrättelsereglementets femgradiga skalor bör användas. Även om deltagarna inte var överens om vilka kategorier eller skalor för osäkerhetsbedömning som är mest lämpliga så upplevde de lösningen i PARC ACH, med kolumner intill evidenslistan, som ett bra sätt att kunna värdera osäkerhet för varje evidens.

Flera deltagare efterfrågade också stöd för att kunna ge ”evidens för evidens”, dvs. att kunna lägga in information som stärker eller försvagar andra evidens. Flera deltagare efterfrågade också möjlighet att kunna gruppera evidens som på något sätt hör ihop.

Samtliga deltagare diskuterade hur avgörande konstruktion och formulering av evidensen är för att på ett bra sätt kunna värdera evidensset. Dessutom behöver analytikern kontrollera att evidensen har olika källor, så att inte samma information förekommer flera gånger och snedvrider analysen.

Två par under det första användartestet ansåg att den nuvarande skalan med två skalsteg för att värdera evidens gentemot det morfologiska fältets tillstånd som antingen stödjande eller motsägande behöver utökas så att det är möjligt att gradera i mer eller mindre stödjande/motsägande. De kunde däremot inte svara på exakt hur många skalsteg som behövdes. Ett annat par ansåg att Hyppos styrka är dess enkelhet och genom att utöka antalet skalsteg blir analysen mer komplicerad utan att för den skull bli mer rättvisande. Deltagarna i användartest 2 lutade åt att några fler skalsteg kan vara relevanta men inte för många. PARC ACH graderar i två steg för stödjande och två steg för motsägande, vilket upplevdes som funktionellt. En deltagare uttryckte att ”systemet ska stödja vår tankeverksamhet men man får inte ge intryck av att systemet kommer att tala om sanningen”. En annan deltagare uttryckte det som att ”för varje ökad komplexitet bygger man på med subjektivitet”.

4.1.3 Tidseffektivitet och kvalitet

Alla deltagarna ansåg att det är för tidskrävande att exkludera tillstånd i den nuvarande versionen. Förutom en enklare hantering med färre knapptryckningar för varje exkludering behövs också en möjlighet att kunna exkludera flera tillstånd samtidigt. Deltagarna var eniga om att bortsett från exkluderingsfunktionen är det inte Hyppos funktionalitet som avgör hur lång tid analysen tar. Det som tar tid är kunskapsprocessen och analysprocessen i form av formulering av evidens, beslut och strukturering av problemet. Hyppo tvingar analytikern att ta ställning och tänka över kopplingar och formuleringar, vilket är tidskrävande men viktigt.

Efterenkäten liksom diskussionerna under användartesterna visar att det var en låg överensstämmelse mellan Hyppos förslag på lösning (hypotes med högst poäng) och den hypotes som deltagarna ansåg som troligast (Enkät: M=2,6, 1=mycket dåligt, 5=mycket bra). Deltagarna såg två huvudorsaker till detta: (1) att de inte hade tid att göra alla exkluderingar av tillstånd vilket innebar att irrelevanta och logiskt motstridiga hypoteser kunde ligga högt upp i hypoteslistan och (2) att det i testscenarierna fanns många stödjande evidens med låg relevans som fick för stor effekt på vilken hypotes som rankades högst. Om flera evidens stödjer (eller motsäger) ett tillstånd så får alla hypoteser som innehåller detta tillstånd en hög (alternativt låg) poäng även om det finns få men starka evidens som pekar i motsatt riktning. För en grupp hamnade också en hypotes som inte hade några evidens kopplade till något av sina tillstånd högt upp i hypoteslistan, vilket ansågs förbryllande. Även detta är en effekt av den mycket enkla algoritmen i Hyppo för att beräkna stödet för en hypotes.

Två par i användartest 1 försökte arbeta enligt ACH:s rekommendation att fokusera på falsifiering snarare än bekräftelse av hypoteser, men hade svårt att hålla fast vid detta. Deltagarna i användartest 2 var eniga om att det bör finnas möjlighet att basera hypotesvärderingen på endast motsägende evidens (idag beräknas hypotesvärderingen i Hyppo utifrån skillnaden mellan stödjande och motsägende evidens). De här deltagarna hade tidigare fått prova på verktyget PARC ACH som enbart summerar motsägende evidens.

Deltagarna ansåg att Heatmap gav ett gott stöd för att bedöma hypoteser (M=4,2, 1=instämmer inte alls, 5=instämmer helt) även om de inte använde den i särskilt stor omfattning under användartesten.

4.1.4 Dokumentation

Samtliga deltagare efterfrågade möjligheten att kunna kommentera kopplingar mellan evidens och tillstånd, uteslutningar, samt hypoteser. Deltagarna vill också kunna se vilka kopplingar, uteslutningar och hypoteser som har en kommentar. Vidare sågs den morfologiska tabellens tillstånd som ett bra redskap för att identifiera informationsbehov. De önskade därför ett bättre sätt att dokumentera informationsbehoven.

Deltagarna efterfrågade också en exportfunktion för att enkelt kunna förmedla och rapportera en analys till någon annan, till exempel genom att exportera analysen till Word. Några deltagare efterfrågade även en namnstämpel så att det är möjligt att se vem som har arbetat med analysen och vad som är gjort.

4.2 Utvärdering av Hyppos användargränssnitt

Deltagarnas svar på enkätfrågorna och kommentarer angående Hyppos gränssnitt och funktionalitet var i stort sett enbart positiva. Deltagarna ansåg att det är lätt

att förstå hur verktyget är tänkt att användas (M= 4,25), att det är enkelt att lägga till nya evidens (M= 4,8) och tillstånd (M=4,5), att de kunde sortera evidenslistan som önskat (M=4,5), att de kunde mata in information i den ordning de ville (M=4,5) och att de förstod vilken information som är möjlig att mata in (M=4,5) (1=instämmer inte alls, 5=instämmer helt).

Deltagarna uppskattade att det är enkelt att flytta runt dimensioner och tillstånd i det morfologiska fältet under analysen. Deltagarna poängterades att det är viktigt att kunna börja analysen var som helst och hoppa mellan rutorna som i den nuvarande versionen, dvs. arbetet ska vara iterativt och inte följa en viss arbetsgång.

Som nämns i avsnitt 5.2.1 efterfrågade samtliga deltagare ett enklare sätt att exkludera tillstånd eftersom det tar för lång tid i den nuvarande versionen.

Vidare menade ett av paren att parenteserna vid tillstånden som visar antalet stödjande och motsägende evidens bör döljas och istället presenteras genom att markera en kryssruta såsom för Heatmap. Det är annars risk för snedvridning av evidensvärderingen. Slutligen påpekade ett annat par att knapparna ovanför det morfologiska fältet (för att värdera evidens som positiva, negativa eller neutrala samt exkluderingsknappen) bör synas även när evidens inte är valda, för att vyerna ska se likadana ut oberoende av vilken vy som man befinner sig i.

4.3 Övriga idéer

Några möjliga intressenter för fortsatta användartester av Hyppo är Länskriminalpolisen och Försvarsmaktens rörliga operativa ledning (ROL). Det kan ge information om hur Hyppo passar för kriminalfall respektive underrättelseanalyser och andra analyser som görs vid operationsplanering.

En återkommande diskussion var hur Hyppo kan stödja identifiering av informationsbehov. Till exempel indikerar tillstånd i det morfologiska fältet som saknar kopplingar till evidens möjliga informationsbehov.

De deltagare som hade erfarenhet av att använda Casper, vilket är den programvara som FOI använder för morfologisk analys, ansåg att Hyppo eventuellt skulle kunna ersätta Casper som inte har någon evidensvärdering i förhållande till tillstånden. Casper har inte heller funktionalitet för att jämföra hypoteser, men däremot funktionalitet för att skapa scenarier som innehåller flera tillstånd för samma dimension, något som saknas i Hyppo. En deltagare efterfrågade den funktionaliteten även i Hyppo.

Slutligen nämnde en deltagare att det kan vara intressant studera om hypoteserna kan värderas med Bayesianska nätverk.

5 Diskussion

Användartesterna visar att Hyppo mycket väl kan användas för att analysera komplexa verkliga analysproblem. Syftet att testa Hyppos funktionalitet kan sägas vara uppnått med råge, då slutsatser kan dras inte bara om användargränssnitt och funktioner utan även avseende konceptet med att kombinera de två analytiska metoderna ACH och morfologisk analys med hjälp av Hyppo.

Samtliga deltagare var överlag positiva till verktygets funktion och tyckte att det var enkelt att använda. Trots den nuvarande prestandabegränsningen som innebär att det morfologiska fältet bör vara max ca 7×5 rutor stort så gör Hyppo det möjligt att hantera många tänkbara hypoteser och att värdera evidens i förhållande till ett morfologiskt fält på ett sätt som inte i något annat verktyg. Resultaten är därför lovande, men användartesterna visar också på ett flertal utvecklingsbehov som behöver hanteras för att Hyppo ska bli mer användbart för professionella analytiker.

De åtta utvecklingsbehov som projektgruppen utifrån användarstudien har identifierat som mest kritiska på kort sikt är att:

- 1) *Använda osäkerheter i analysen:* Inom Försvarsmakten används till exempel sakriktighet, tillförlitlighet och relevans på evidens som mått på osäkerheter som behöver vägas in i analysen. Ett möjligt sätt är att införa möjligheter att skatta dessa och förändra beräkningsalgoritmen så att osäkra eller mindre relevanta evidens ges mindre vikt i analysen. Ett annat är att utesluta evidens med för hög osäkerhet. Det sätt som verktyget PARC ACH använder trovärdighet och relevans på upplevdes som enkelt att använda av deltagare i användartest 2.
- 2) *En rättvisande hypotesranking:* Hyppos ranking av hypoteser baseras på den ”poäng” som automatiskt beräknas för varje hypotes. I nuläget ger alla evidens som stödjer ett tillstånd i en viss hypotes värdet +1 och alla motsägande evidens värdet -1 i sammanräkningen av hur mycket stöd den aktuella hypotesen har. Förhållandet mellan + och - avgör alltså hypotesens poäng. En fördel med denna enkla poängberäkningsalgoritm är att metoden blir lätt att förstå. Nackdelen är att enkelheten leder till en del oönskade effekter. Exempelvis så kan en intuitivt osannolik hypotes med tillstånd som helt saknar värdering hamna högt upp på rankinglistan. En beslutsstödsfunktion som varnar när det förekommer tillstånd som inte har kopplats till något evidens skulle kunna vara en framtida funktion för att förbättra hypotesvärderingen. Att vikta evidens med avseende på deras relevans och trovärdighet (enl. punkten ovan) kan också leda till mer rättvisande analyser. Till exempel uppstår det ibland situationer där många mindre relevanta evidens pekar i en

riktning samtidigt som något enstaka men viktigt evidens pekar i en annan riktning. I sådana situationer får de mindre relevanta evidensen orimligt stor vikt med den poängberäkning som används i Hyppo idag. Resultatet även i detta fall kan bli att hypoteser som bedöms som osannolika av en analytiker rankas högt av Hyppo och vice versa. Samtidigt som mycket talar för att hypotesrankingen i Hyppo behöver förbättras är det viktigt att den inte görs alltför avancerad. Analytikern ska inte få en övertro till systemets värdering, vilket skulle kunna vara en risk om värderingen görs mer avancerad.

- 3) *Möjliggöra koppling av evidens till kombinationer av tillstånd:* Försöksdeltagarna i användarförsöken upplevde det som ett problem att evidens endast kunde stödja och/eller motsäga enskilda tillstånd i det morfologiska fältet. Att inte kunna koppla evidens till kombinationer av tillstånd upplevdes som begränsande och det påverkade också kvaliteten på hypotesrankingen märkbart.
- 4) *Skriva in informationsbehov och kommentera:* Olika delar i analysen behandlas successivt som till exempel exkluderingar, kopplingar mellan tillstånd och evidens, samt hypoteser. Deltagarna upplevde ett behov av att dokumentera sina beslut på ett bättre sätt eftersom de iterativt återkommer till samma delar för att utforma det morfologiska fältet och strukturera evidens.
- 5) *Förenkla proceduren för exkluderingar:* Den nuvarande implementationen upplevde deltagarna i testet som krånglig och tidskrävande.
- 6) *Förbättra prestandan:* I den nuvarande versionen så hänger sig Hyppo om det morfologiska fältet blir för stort. Det måste gå att använda ett morfologiskt fält med till exempel tio dimensioner och tio tillstånd i varje dimension utan att Hyppo blir långsamt.
- 7) *Förbättra export- och importfunktioner:* Att lägga in beskrivningar av evidens, dimensioner, tillstånd m.m. i Hyppo är tidskrävande eftersom det görs manuellt. Dessa beskrivningar, tillsammans med resultaten av analysen behöver kunna exporteras på ett format som kan användas för rapportering.
- 8) *Förbättra versionshanteringen och spåra ändringar:* För att flera analytiker ska kunna arbeta tillsammans på ett problem över lång tid så behövs funktioner för att indikera vem som har varit inne i analysen och loggning av vilka förändringar som gjorts. Analysen måste låsas så att inte flera personer kan ändra analysen samtidigt och skriva över varandras ändringar.

Förslag på hur några av dessa funktioner skulle kunna utformas i Hyppo finns beskrivna i exjobbssrapporten ”Hur analytiska beslutsstödsystem kan stödja underrättelsetjänstens arbete inom den svenska försvarsmakten” (von Feilitzen, 2014). I intervjuerna för användartest 2 fick försöksdeltagarna frågan om huruvida de ofta jobbar med en viss typ av analysproblem, och om det i så fall skulle vara intressant att på något sätt stödja återanvändning av deras gamla analyser i Hyppo. Samtliga försöksdeltagare uppgav dock att de inte har några ”typiska” uppdrag och att det därför skulle vara svårt att återanvända data på det sätt som föreslogs.

Värt att notera är att deltagarna uppskattade Hyppos enkelhet och den riskerar att gå förlorad om det blir för många komplicerade funktioner. Om man till exempel gör det möjligt att koppla evidens till kombinationer av tillstånd så blir evidensvärderingen mer komplicerad. Man får möjlighet att göra en bättre och mer rättvisande evidensvärdering, men risken för att analytikern själv ska råka införa logiska fel i analysen ökar. Om det till exempel finns evidens för att om snickaren begick mordet så måste det ha skett med hammaren så är det lätt att analytikern kopplar ihop gärningsmannen ”snickaren” med mordvapnet ”hammaren”. Detta innebär dock att analytikern har gjort ett logiskt fel. Evidens som motsäger ”snickaren” behöver inte betyda att ”hammaren” ska ges lägre poäng eftersom det kan finnas andra än snickaren som hade tillgång till hammaren.

Ett utvecklingsbehov på längre sikt, som förmodligen inte kommer att kunna lösas inom ramen för det pågående projekt, är att möjliggöra skapandet av scenarier som innehåller fler än ett tillstånd från varje dimension. Hyppo är inte utvecklat för att kunna hantera denna typ av problem, men om Hyppo ska kunna användas även för traditionell morfologisk analys behöver denna funktionalitet stödjas. En annan funktionalitet som finns i Casper men inte i Hyppo är möjligheten att skapa undermodeller när det morfologiska fältet blir för komplext. När undermodeller skapas är det också viktigt att evidensen följer med.

Det kan också finnas behov av att på enklare sätt än idag kunna skapa en kopia av analysen för att kunna pröva olika sätt att omstrukturera det morfologiska fältet. I Hyppo är det möjligt att välja ”spara som” med nytt filnamn men det blir några extra knapptryckningar för att växla mellan olika analyser. I PARC ACH går det att skapa kopior av en analys som listas i en ruta så att det är enkelt att växla mellan originalmodellen och de alternativa modellerna samt välja vilka modeller som ska sparas och vilka som ska raderas.

6 Slutsats

Med hjälp av Hyppo är det möjligt att koppla ihop två analytiska metoder på ett sätt som ger nya möjligheter som de båda metoderna var för sig inte klarar av. Avseende den morfologiska analysen tillförs dimensionen att kunna koppla evidens till det morfologiska fältets tillstånd, vilket innebär att det är möjligt att värdera kombinationer av tillstånd på ett sätt som inte görs på ett lika tydligt sätt i traditionell morfologisk analys. Jämfört med traditionell ACH finns möjlighet att med hjälp av Hyppo hantera fler och mer komplexa hypoteser. De analytiker som medverkade under försöken hade enbart positiva synpunkter avseende att arbeta enligt den analysmetod som Hyppo stödjer.

För att få riktigt hög kvalitet i analyser med Hyppo krävs fortsatt utveckling avseende framför allt osäkerhets- och hypotesvärdering, möjlighet till dokumentation samt export av analyser för att underlätta delat arbete och rapportering. Användartesterna visade dock tydligt att verktyget redan idag ger ett mervärde för genomförandet av analyser och att gränssnittet upplevs som tilltalande och enkelt.

7 Referenser

Försvarsmakten. (2010). *Försvarsmaktens Underrättelsereglemente (FM UndR)*. M7739-353025. Stockholm: Försvarsmakten

Forsgren, R., Kaati, L., Mårtenson, C. Svenson, P., & Tjörnhammar. E. (2008). An overview of the impactorium tools 2008. I *Proceedings of the Second Skövde Workshop on Information Fusion Topics (SWIFT 2008)*. 4-6 November, Skövde.

Carlsen, H. & Dreborg, K.H (2008)., *Dynamisk generering av socio-ekonomiska scenarier för klimatanpassning: metod, byggstenar och exempel*. FOI-R--2512--SE

Globalytica (2014). *TH!NK Suite*. Globalytica, LLC.

Gustavi, T., Johansson, F., Johansson, R., Jändel, M., Kaati, L., Mårtenson, C., & Oskarsson, D. (2011). *Analytiska verktyg - en översikt av metoder och modeller för underrättelseanalys*. FOI-R--3310—SE. Stockholm: Totalförsvarets forskningsinstitut.

Gustavi, T., Johansson, F., Mårtenson, C., & Svenson, P. (2011). *Datorstöd för hypoteshantering inom underrättelseanalys*. FOI-R--3214--SE. Stockholm: Totalförsvarets forskningsinstitut.

Gustavi, T., Karasalo, M., & Mårtenson, C. (2013). A tool for generating, structuring, and analyzing multiple hypotheses in intelligence work. I *Proceedings of the 2013 European Intelligence and Security Informatics Conference (EISIC 2013)* (pp. 23-30). 12-14 Augusti 2013, Uppsala, Sverige. IEEE.

Heuer, R. J. (1999). *Psychology of Intelligence Analysis*. Washington, DC: CIA Center for the Study of Intelligence.

Jones, M. D. (1998). *The Thinker's Toolkit: 14 Powerful Techniques for Problem Solving*. New York: Three Rivers Press.

Lehner, P.E., Adelman, L., Cheikes, B.A., & Brown, M.J. (2008). Confirmation bias in complex analyses. *IEEE Transactions on System, Man and Cybernetics, Part A: Systems and Humans*, 38(3), 584-592.

PARC (2010). ACH. Palo Alto, CA: Palo Alto Research Center.

Pirolli, P. L., & Card, S. (2005). The sensemaking processes and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of the 2005 International Conference on Intelligence Analysis*, May 2-5, McLean, VA.

Ritchey, T. (2006). Problem structuring using computer-aided Morphological Analysis. *Journal of the Operational Research Society*, 57, 792–801.

Roy, J. & Guyard, A. B. (2010). Multiple hypothesis situation analysis support system prototype. In *Proceedings of the 13th Conference on Information Fusion (FUSION 2010)* (pp. 1-8). 26-29 juli, 2010, Edinburgh, IEEE.

Stenström, M. (2013). *Morphological Analysis in Groups: A Personal Guide*. FOI-R--3678--SE. Stockholm: Totalförsvarets forskningsinstitut.

von Feilitzen, J. (2014). Hur beslutsstödsystem för analytiskt arbete kan stödja underrättelsetjänstens arbete inom den svenska Försvarsmakten. Examensarbete i Management, kommunikation och IT vid Uppsala Universitet, 2014.

Zwicky, F. (1969). *Discovery, Invention, Research through Morphological Analysis*. New York: McMillan

Bilaga 1. Utvecklingsbehov

Tabell 1 till 5 sammanfattar utvecklingsbehoven från användartesterna. Den första kolumnen beskriver behovet och den andra kolumnen innehåller kommentarer från projektgruppen om till exempel genomförbarhet eller behov av kompletterande utredning. I samtliga tabeller är behov som identifierades av flera grupper markerade med *.

Tabell 1. Utvecklingsbehov för gränssnittet.

Utvecklingsbehov	Kommentar och lösningsförslag
Att nya evidens inte försvinner in i listan, vilket nu blir fallet om det är valt att evidens ska sorteras enl. bokstavsordning*	
Excel-liknande kolumnhantering	Exempelvis automatisk radbrytning av textinformationen i rutorna när dessa minskas, samt möjlighet att ändra kolumnbredd.
Fler än två tillstånd kan markeras för exkludering men bara två av dem exkluderas	
Knapparna för värdering ska visas oavsett vilken vy som visas	Förenklar användningen.
Kunna dölja informationen om antalet stödjande och motsägande tillstånd	Eventuellt med kryssruta som för Heatmap.
Kunna markera evidens och se vilka hypoteser som berörs	
Kunna ta bort ett tillstånd genom att trycka på delete istället för vänsterklick*	Förenklar användningen. Eventuellt behövs en varning för att bekräfta borttagningen. Eventuellt behövs samma möjlighet för att ta bort evidens.

Tabell 2. Utvecklingsbehov för dokumentation.

Utvecklingsbehov	Kommentar och lösningsförslag
Dokumentera informationsbehov*	Placering?
Stämpla analysen så att det är möjligt att se vem som har gjort ändringar i analysen	Behövs om flera analytiker arbetar med analysen vid olika tillfällen.
Kunna logga vilka förändringar som har gjorts för en viss version av analysen	
Skriva kommentarer och anteckningar, exempelvis att kommentera de kopplingar som görs mellan evidens och tillstånd*	Markeringar behövs som visar var det finns kommentarer
Kunna exportera analysen till ett vanligt format som till exempel Word*	Det oklart hur alla vyer och kopplingar inklusive kommentarer ska beskrivas i ett dokument.

Tabell 3. Utvecklingsbehov för evidenslista och evidensvärdering.

Utvecklingsbehov	Kommentar och lösningsförslag
Kunna ange osäkerhet/kvalitetsindikator på evidens*	Sakriktighet och tillförlitlighet kan anges enligt FM UndR men även enligt andra kategorier. Ett annat alternativ är att ge analytikern möjlighet att välja kategorier. Osäkerheterna kan till exempel anges med ytterligare kolumner i evidenslistan. Det är oklart hur osäkerheterna ska värderas i analysen.
Kunna välja om kvalitetsindikatorer ska räknas med/viktas in i analysen	Se PARC ACH.
Fler skalsteg för värdering av evidens mot tillstånd*	Grupperna var oeniga. Det är också oklart om det ska vara möjligt att välja neutralt och/eller not applicable, samt välja antal skalsteg.
Kunna gruppera evidens	Gruppering kan ge en bättre visuell överblick. Det är även möjligt att värderingsberäkningarna kan påverkas.

Förbättrad indikering av evidens som inte har hanterats	Ändringar i det morfologiska fältet gör ofta att alla evidens indikeras för ytterligare hantering trots att ändringen inte påverkar evidensvärderingen. Eventuellt bör analytikern kunna ta bort indikeringen manuellt.
Ange evidens för evidens, dvs. information som gör evidens mer eller mindre tillförlitliga*	Vissa evidens är mer relaterade till andra evidens. Det är även möjligt att värderingsberäkningarna kan påverkas.
Kunna lägga in länk/bifoga artikel som stödjer evidens	
Kunna värdera evidens mot kombinationer av tillstånd*	Det är oklart hur kombinationerna ska användas i värderingsberäkningarna.
Kunna välja om positiva evidens ska användas i hypotesvärderingen	PARC ACH använder enbart negativa evidens.
Kunna skilja på objektiva evidens och antaganden	Eventuellt ska det gå att välja om antagandena ska användas i analysen.
Förbättra hypotesvärderingen så att inte mindre relevanta evidens får för stor effekt*	Det är oklart hur det faktiska informationsvärdet ska beskrivas för flera evidens som delvis säger samma sak.
Förbättra visualiseringen av hypoteser som har få eller inga kopplade evidens	Även hypoteser helt utan kopplade evidens kan komma högt upp i listan.

Tabell 4. *Utvecklingsbehov för morfologiskt fält.*

Utvecklingsbehov	Kommentar och lösningsförslag
Förenkla exkludering av tillståndskombinationer*	Till exempel genom att klicka i det morfologiska fältet och/eller i hypoteslistan.
Kunna koppla exkluderingar till antaganden, evidens eller logik*	Till exempel genom kommentarer eller kopplingar till antaganden eller evidens.
Kunna exkludera flera tillstånd på en gång.	Det måste vara lätt att förstå vilka kombinationer som exkluderas.
Flera grader av exkludering så att osannolika men inte omöjliga hypoteser finns kvar i hypoteslistan men flyttas långt ned*	Det är oklart hur många grader av osannolika hypoteser som behövs och hur osannolika hypoteser ska presenteras.
Kunna kombinera tillstånd i olika dimensioner för gemensam evidensvärdering*	Det är oklart hur evidensvärderingen av enskilda tillstånd i kombinationen ska påverka värderingen av övriga tillstånd. Att ge samma värdering till de övriga tillstånden riskerar att ge logiska fel. Det kan vara bättre att sammanfoga tillstånden till ett nytt tillstånd.
Kunna skapa undermodeller	Det är oklart hur koppling mellan undermodell och modermodell ska realiseras.
Kunna skapa flera alternativa morfologiska fält utifrån samma evidenslista	Se PARC ACH.
Kunna skapa scenarier bestående av flera tillstånd ur varje dimension	Se Casper. Funktionen är efterfrågad vid morfologiska analyser.

Tabell 5. *Övriga utvecklingsbehov.*

Utvecklingsbehov	Kommentar och lösningsförslag
Öka prestandan så att systemet klarar ett morfologiskt fält med minst tio dimensioner och tio tillstånd i varje dimension	Till exempel genom att analytikern anger när hypotesvärderingen ska göras och förbättrad programkod.
Kunna visualisera hypoteser i en 2D-karta eller tid och rum, samt gruppera hypoteser	

FOI är en huvudsakligen uppdragsfinansierad myndighet under Försvarsdepartementet. Kärnverksamheten är forskning, metod- och teknikutveckling till nytta för försvar och säkerhet. Organisationen har cirka 1000 anställda varav ungefär 800 är forskare. Detta gör organisationen till Sveriges största forskningsinstitut. FOI ger kunderna tillgång till ledande expertis inom ett stort antal tillämpningsområden såsom säkerhetspolitiska studier och analyser inom försvar och säkerhet, bedömning av olika typer av hot, system för ledning och hantering av kriser, skydd mot och hantering av farliga ämnen, IT-säkerhet och nya sensorers möjligheter.



FOI
Totalförsvarets forskningsinstitut
164 90 Stockholm

Tel: 08-55 50 30 00
Fax: 08-55 50 31 00

www.foi.se